


```

          447 GRGGAGCTACGGTGTACTTAACACTATCCACAGCTTGCGGTTTAGCT   496
      151 GYAlAIAIaIA 154
            |
            |
            |
          497 GGTCGCCGCAGCT 508

seq_name: gb_hlg:AP003507

seq_documentation_block:
LOCUS       AP003507             149040 bp           DNA               12-Apr-2001
DEFINITION  Oryza sativa chromosome 6 clone P0468A12, *** SEQUENCING IN PROGRESS ***, In ordered pieces.
ACCESSION   AP003507
VERSION     AP003507.1    GI:13603466
KEYWORDS    HG; HTGS_PHASE2.
SOURCE      Oryza sativa (cultivar:Nipponbare) DNA, clone:P0468A12.
ORGANISM    Oryza sativa
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoidae; Oryzaceae; Oryza.

REFERENCE   1 (sites)
AUTHORS     Sasaki,T., Matsumoto,T. and Yamamoto,K.
TITLE       Oryza sativa nipponbare(Gx3) genomic DNA, chromosome 6, PAC clone:P0468A12
COMMENT     Published Only in DataBase (2001) In press
JOURNAL     2 (bases 1 to 149040)
AUTHORS     Sasaki,T., Matsumoto,T. and Yamamoto,K.
TITLE       Direct Submission
REFERENCE   Submitted (11-APR-2001) Takuji Sasaki, National Institute of Agricultural Resources, Rice Genome Research Program; Kannondai 2-1-2, Tsukuba, Ibaraki 305-8602, Japan
              (E-mail:tsusakia@affrc.go.jp, URL:http://rpg.dna.affrc.go.jp/, Tel:81-298-38-7441, Fax:81-298-38-7468)
COMMENT     NOTE: It currently consists of 1 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have provided by the submitter. This sequence will be replaced by the finished sequence as soon as it is available and the accession number will be preserved.
              * NOTE: This is a "working draft" sequence.
              * This sequence will be replaced
              * by the finished sequence as soon as it is available and
              * the accession number will be preserved.

FEATURES             source
         location/Qualifiers
             1..149040
                /organism="Oryza sativa"
                /cultivar="Nipponbare"
                /db_xref="taxon:4530"
                /chromosome="6"
                /clone="P0468A12"

BASE COUNT  41830 a 33432 c 32383 g 40945 t 450 others

ORIGIN
alignment_scores:
Quality:        220.00          Length:        129
Ratio:          3.099           Gaps:        6
Percent Similarity: 55.039      Percent Identity: 41.085

alignment_block:
US-09-554-547-16 x AP003507/rev ..

Align seg 1/1 to reverse of: AP003507 from: 1 to: 149040

17 AsnAtaAatgAlaGlYargLeuglSerAspleuAspThPhneGlyar 33
|||||:::||||| ||| ::|::|||
81412 AATGTGAAGCCAAAGCAGTCAAGGGCTCAGAGCATCACAGTACGAGCGG 81363
33 gvaHicISglYanleuTYrAlaGlyTlGeluaGrAlaGlyProArvGeLYT 50
:::::::::::||||
81362 TGCTATTGAGAGAGATAT.....GGAT 81340
```

```

50 yrPrroGlyLeuThrAlaSerIleGlyGluValGlyAlaArgLeuGly 66
|||
81339 AT.....GGC 81335
67 GLyArGrAGlYvAlGlyAlSeRSeTyrgLYTYrGLyTYrPProSeTr 83
|||||
81334 GGCGGGGTATGGTGAGGA.....GGATATGGTCATCTGGGCTA 81297
83 pGLyTYrPProTYrGLyGLyTYrGLyGLyTYrGLyGLyTYrGLyGLy 98
:||| ||||| |||||
81296 TGcTGTGcTGTgATGcTGgATATGcCCACcCTGcATATGcGGGcGCT 81247
99 ..TYrGLyTYrAspGlnGlnPhGeLySeRlaTYrGLyGLyTYrPro 114
|||||
81246 ACcGAGcGTGcATAcGCCCAcAGATATGcCTGTGGcTAT...GGTCATcCT 81200
115 GLyTYrTYrGLyTYrTYrTYrTYrPProSeRGlyTYrGLyGLyTYrGLy 131
|||||
81199 GGTCACAGTGA.....GGATATGGTGcTGTtACcGAGc 81165
131 ySeRTyrGLyGLySeRTyrGLyGLySeRTyrTYrTYr 143
|||||
81164 TGGTTATGGTGAGATAcGcTGTGGGcGTGcCTAc 81128
seq_name: gb_in:DROEDG91A
seq_documentation_block:
LOCUS DROEDG91A 1940 bp DNA INV 26-APR-1993
DEFINITION Drosophila melanogaster EDG-91 gene, complete cds.
ACCESSION M71250
VERSION M71250.1 GI:157326
KEYWORDS cuticle protein.
SOURCE Drosophila melanogaster DNA.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
AUTHORS Apple,R.T. and Fristrom,J.W.
TITLE 20-hydroxyecdysone is required for, and negatively regulates,
transcription of Drosophila pupal cuticle protein genes
JOURNAL Dev. Biol. 146, 569-582 (1991)
MEDLINE 91323677
FEATURES
source location/Qualifiers
1..1940
/oranism="Drosophila melanogaster"
/db_xref="taxon:7227"
1053..1059
/gene="EDG-91"
join(1118..1141,1205..1660)
/gene="EDG-91"
join(1118..1141,1205..1660)
/gene="EDG-91"
/note="cuticle protein"
/codon_start=1
/db_xref="GI:157327"
/translation="MALVRSQMLALLIAGOGAAPKTKTGRLLGLGFGSGSVIGI
SAGIVGGGILYSFGGGGYPGVASGYPGVGGSYNVNGSGSGFSGGYPGGYSYG
FGHPRHHGGYSGGSGSYHNDGSGYGGYSOSQYSNQYGGGSGGSGGNCFGSK
<1118..1141
/gene="EDG-91"
/number=1
1205..1738
/gene="EDG-91"
/number=2
1723..1729
/gene="EDG-91"
1739
/gene="EDG-91"
exon
polyA_site
BASE COUNT 516 a 365 c 468 g 591 t
ORIGIN
```

alignment_scores: Quality: 219.00 Length: 172
 Ratio: 2.235 Gaps: 11
Percent Similarity: 56.977 Percent Identity: 38.372

Alignment_block:
US-09-554-547-16 x DROEDG91A ..

Align seg 1/1 to: DROEDG91A from: 1 to: 1940

```

7 LeuSerLeuLeuSerThrAlaAlaLeuThrAsnAlaAlaArgAlaGlyAr 23
   ::::::::::::::::::::: :::::::::::::::
1118 ATGCGTCTGTCGCGGTGAGTGTCTGTAAGTCGCGCTGCTATTTCCGCTCC 1167
   ::::::::::::::::::::: :::::::::::::::
23 gLeuGlySer.....AspLeuAspThrPheGlyArGValHisGlyA 37
   ::::::::::::::::::::: :::::::::::::::
1168 GATTCGGATGCACATCAATCGATTTGTTACCTTCGACAGATCGTCCCTTT 1217
   ::::::::::::::::::::: :::::::::::::::
37 snLeuTyAlaGlyIleGluArgAlaGlyProArgGlyTyProGlyLeu 53
   ::::::::::::::::::::: :::::::::::::::
1218 TCCTGATTTCGCCGTCAAGTCACAGCGCGCCGCGTGAAGACCAAGGTGCG 1267
   ::::::::::::::::::::: :::::::::::::::
54 ThrAlaSer...IleGlyGlyIleValGlyAlaArgLeuGlyIleArgAl 69
   ::::::::::::::::::::: :::::::::::::::
1268 ACCTTGGCGCTTCGTGGCGGTGATTTGTGTGACAGTACGACTTATGTC 1317
   ::::::::::::::::::::: :::::::::::::::
69 aGlyValIleValIleSerSer.....TyrGlyTyArgIleTy 80
   ::::::::::::::::::::: :::::::::::::::
1318 CGGCATTCGACGCTGCGTGTGCTGTATACGCGTTTCGAGCGCGGTGCT 1367
   ::::::::::::::::::::: :::::::::::::::
80 yrPro.....SerTrpGlyTyProTyArgIleTyArgIleTyArgIle 92
   ::::::::::::::::::::: :::::::::::::::
1368 ATCTCGTGTGCGCTATGCGAGTGTGATACCCA...GTTGATATGTGTGTGCG 1414
   ::::::::::::::::::::: :::::::::::::::
93 TyrGlyGlyIleTyArgIleTyArgIleTyArgIleTyArgIleTyArgIle 109
   ::::::::::::::::::::: :::::::::::::::
1415 TACTCAGCGCTATACGCGCTACGAGAGCG.....AGTGGATTGCGAGTGTG 1458
   ::::::::::::::::::::: :::::::::::::::
109 aTyArgIleTyArgIleTyArgIleTyArgIleTyArgIleTyArgIle 117
   ::::::::::::::::::::: :::::::::::::::
1459 CTACTATCCAGAGAGAGGTGTACTCCGCGTTTGGACACAGCGCGCATTTAC 1508
   ::::::::::::::::::::: :::::::::::::::
117 yrGlyTyArgIleTyArgIleTyArgIleTyArgIleTyArgIleTyArgIle 128
   ::::::::::::::::::::: :::::::::::::::
1509 ACGGAGATACATATCCGGCGGTGATGTACCAACAATCAGGCGCGGATCT 1558
   ::::::::::::::::::::: :::::::::::::::
129 TyrGlyGlySerTyArgIleTyArgIleTyArgIleTyArgIleTyArgIle 134
   ::::::::::::::::::::: :::::::::::::::
1559 TATGCGCGCGCACTATATGTCAGTCACTACGATGATGATTATTACGAGAG 1608
   ::::::::::::::::::::: :::::::::::::::
134 yGlySerTyArgIleTyArgIleTyArgIleTyArgIleTyArgIleTyArgIle 139
   ::::::::::::::::::::: :::::::::::::::
1609 TGGTGGCTATGAGAGCG 1624

```

seq_name: gb_hlg:AC017979

seq_documentation_block:
LOCUS AC017979 23547 bp DNA 09-DEC-1999
DEFINITION Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***, in ordered
pieces.
ACCESSION AC017979
VERSION AC017979.1 GI:6553211
KEYWORDS HMG; HMGs_PHASE2.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
1 (bases 1 to 23547)
AUTHORS Adams, M. and Venter, J.C.
TITLE Direct Submission
JOURNAL Submitted (09-DEC-1999) Celera Genomics, 45 West Gate Drive,

COMMENT
Rockville, MD, USA
This sequence was identified as CDW:10212867 by the submitter.
For more information on this record e-mail to fly@celera.com.
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

FEATURES
source
1. 23547
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"

BASE COUNT 6844 a 4896 c 4767 g 7040 t
ORIGIN

alignment_scores: Quality: 215.00 Length: 172
 Ratio: 2.194 Gaps: 11
Percent Similarity: 56.977 Percent Identity: 37.791

Alignment_block:
US-09-554-547-16 x AC017979 ..

Align seg 1/1 to: AC017979 from: 1 to: 23547

```

7 LeuSerLeuLeuSerThrAlaAlaLeuThrAsnAlaAlaArgAlaGlyAr 23
   ::::::::::::::::::::: :::::::::::::::
19905 ATGCGTCTGTCGCGGTGAGTGTCTGTAAGTCGCGCTGCTATTTCCGCTCC 19954
   ::::::::::::::::::::: :::::::::::::::
23 gLeuGlySer.....AspLeuAspThrPheGlyArGValHisGlyA 37
   ::::::::::::::::::::: :::::::::::::::
19955 GATTCGGATGCACATCAATCGATTTGTTACCTTCGACAGATCGTCCCTTT 20004
   ::::::::::::::::::::: :::::::::::::::
37 snLeuTyAlaGlyIleGluArgAlaGlyProArgGlyTyProGlyLeu 53
   ::::::::::::::::::::: :::::::::::::::
20005 TCCTGATTTCGCCGTCAAGTCACAGCGCGCCGCGTGAAGACCAAGGTGAC 20054
   ::::::::::::::::::::: :::::::::::::::
54 ThrAlaSer...IleGlyGlyIleValGlyAlaArgLeuGlyIleArgAl 69
   ::::::::::::::::::::: :::::::::::::::
20055 ACCTTGGCGCTTCGTGGCGGTGATTTGTGTGACAGTACGACTTATGTC 20104
   ::::::::::::::::::::: :::::::::::::::
69 aGlyValIleValIleSerSer.....TyrGlyTyArgIleTy 80
   ::::::::::::::::::::: :::::::::::::::
20105 CGGCATTCGACGAGTGTGCGTGTATACGCGTTTTCGAGCGCGGTGCT 20154
   ::::::::::::::::::::: :::::::::::::::
80 yrPro.....SerTrpGlyTyProTyArgIleTyArgIleTyArgIleTyArgIle 92
   ::::::::::::::::::::: :::::::::::::::
20155 ATCTCGTGTGCGCTATGCGAGTGTGATACCCA...GTTGATATGTGTGTGCG 20201
   ::::::::::::::::::::: :::::::::::::::
93 TyrGlyGlyIleTyArgIleTyArgIleTyArgIleTyArgIleTyArgIleTyArgIle 109
   ::::::::::::::::::::: :::::::::::::::
20202 TACTCAGCGCTATACGCGCTACGAGAGCG.....AGTGGATTGCGAGTGTG 20245
   ::::::::::::::::::::: :::::::::::::::
109 aTyArgIleTyArgIleTyArgIleTyArgIleTyArgIleTyArgIleTyArgIle 117
   ::::::::::::::::::::: :::::::::::::::
20246 CTACTATCCAGAGAGAGGTGTACTCCGCGTTTGGACACAGCGCGCATTCAC 20295
   ::::::::::::::::::::: :::::::::::::::
117 yrGlyTyArgIleTyArgIleTyArgIleTyArgIleTyArgIleTyArgIleTyArgIle 128
   ::::::::::::::::::::: :::::::::::::::
20296 ACGGAGATACATATCCGGCGGTGATGTACCAACAATCAGGCGCGGATCT 20345
   ::::::::::::::::::::: :::::::::::::::
129 TyrGlyGlySerTyArgIleTyArgIleTyArgIleTyArgIleTyArgIleTyArgIle 134
   ::::::::::::::::::::: :::::::::::::::
20346 TATGCGCGCGCACTATATGTCAGTCACTACGATGATGATATTACGAGAG 20395
   ::::::::::::::::::::: :::::::::::::::
134 yGlySerTyArgIleTyArgIleTyArgIleTyArgIleTyArgIleTyArgIle 139
   ::::::::::::::::::::: :::::::::::::::
20396 TGGTGGCTATGAGAGCG 20411

```

seq_name: gb_in:AC009461

seq_documentation_block:
LOCUS AC009461 171612 bp DNA INV 23-MAR-2001

DEFINITION Drosophila melanogaster, chromosome 3R, region 90B-90C, BAC clone BACR1113, complete sequence.

ACCESSION AC009461

VERSION AC009461.5 GI:13435227

KEYWORDS HTG.

SOURCE fruit fly

ORGANISM Drosophila melanogaster

REFERENCE Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

AUTHORS 1 (bases 1 to 171612)

Celniker, S.E., Adams, M.D., Krommler, B., Tyler, D., Wan, K.H., Holt, R.A., Evans, C.A., Gocayne, J.D., Amanatides, P.G., Brandon, R.C., Rogers, Y., An, H., Baldwin, D., Bazon, J., Beeson, K.Y., Busam, D.A., Carlson, J.W., Center, A., Champs, M., Davenport, L.B., Dietz, S.M., Dodson, K., Dorsett, V., Doup, L.E., Doyle, C., Dresnek, D., Farfan, D., Ferreira, S., Frisoe, E., Galle, R.F., Garg, N.S., George, R.A., Gonzalez, M., Houck, J., Hoskins, R.A., Hostin, D., Howland, T.J., Ibegwam, C., Jalali, M., Kruse, D., Li, P., Mattei, B., Moshrefi, A., McIntosh, T.C., Moy, M., Murphy, B., Nelson, C., Nelson, K.A., Nunoo, J., Pacleb, J., Paragas, V., Park, S., Patel, S., Pfeiffer, B., Phouanavong, S., Pittman, G.S., Puri, V., Richards, S., Scheeler, F., Stapleton, M., Strong, R., Svirskaas, R., Tector, C., Williams, S.M., Zaveril, J.S., Smith, H.O., Rubin, G.M. and Venter, J.C.

TITLE Sequence of Drosophila chromosome 3R, region 90B-90C

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 171612)

AUTHORS Celniker, S.E., Agbayani, A., Arcaina, T.T., Baxter, E., Blazek, R.G., Burenhoff, C., Champs, M., Chavez, M., Chew, M., Ciesiolka, L., Doyle, C.M., Farfan, D.E., Galle, R., George, R.A., Harris, N.L., Hoskins, R.A., Houston, K.A., Hummasti, S.R., Karra, K., Kearney, L., Kim, E., Lee, B., Lewis, S., Li, P., Lomotan, M.A., Mazda, P., Moshrefi, A.R., Moshrefi, M., Nixon, K., Pacleb, J.M., Park, S., Pfeiffer, B., Poon, L., Sequeira, A., Sethi, H., Snir, E., Svirskaas, R.R., Wan, K.H., Weinburg, T., Zhang, R., Zietan, L.L. and Rubin, G.M.

TITLE Direct Submission

JOURNAL Submitted (23-AUG-1999) Drosophila Genome Center, Lawrence Berkeley Laboratory, MS 64-121, Berkeley, CA 94720, USA

COMMENT On Mar 23, 2001 this sequence version replaced gi:5836088.

Sequence submitted by:

Berkeley Drosophila Genome Project

Lawrence Berkeley National Laboratory, MS 64-121

Berkeley, CA 94720

This sequence was assembled using end sequences from a whole genome shotgun and from subclones of this BAC and its neighboring clones. For further information about this sequence, including its location and relationship to other sequences, please visit our sequence archive web site (<http://www.fruitfly.org/sequence/>) or send email to bdg@fruitfly.berkeley.edu.

Location/Qualifiers

1. 171612

/organism="Drosophila melanogaster"

/strain="y; cn bw sp"

/dd_xref="taxon:7227"

/chromosome="3R"

/map="90B-90C"

/clone="BACR1113 (D981)"

/clone_lib="RPCT-98 (Roswell Park Cancer Institute Drosophila melanogaster BAC library, partial EORI in PBACE3.6)"

BASE COUNT 50924 a 36009 c 35092 g 49587 t

ORIGIN

alignment_scores:

Quality: 215.00 Length: 172

Ratio: 2.194 Gaps: 11

Percent Similarity: 56.977 Percent Identity: 37.791

alignment_block:

US-09-554-547-16 x AC009461/rev

Align seg 1/1 to reverse of: AC009461 from: 1 to: 171612

7 LeuSerLeuSerThraLaalaleThraAsnAlaAlaArgAlaGlyAr 23

110970 ATGGCTGTGGTTCGGCTGAGTTGTGTAGTCCGGCTGATTTCCTCC 110921

23 gleuGlySer.....AspleuAspThrPheGlyArGValHisGlyA 37

110920 GATTGGAGTACACTGAAATGATTTGGTACCTTGACATGCGCCCTTT 110871

37 snLeuTrAlaGlyIleGluArGAlaGlyProArgIlyrProGlyLeu 53

110870 TGTGTATTCCTGGTCAAGTCAAGTCGGCCGGTGAAGCCGAAGGTACC 110821

54 ThrAlaSer...IleGlyIlyGluValIlyAlaArgLeuGlyIlyrGAl 69

110820 ACCTTGGCGCTTCTGGCGCGTGGATTTGGTGGCAGCTAGATTTAGTGC 110771

69 aGlyValIlyAlaSerSer.....TyrGlyTyrGlyT 80

110770 CGGCATCGAGTACGTCGTCGCTGTATAGCGGTTTCGAGCGCGTGCCT 110721

80 yPro.....SerTPGlyTyrProTyrGlyIlyr...GlyGly 92

110720 ATCTGTGTGGCTATGCGGAGTGGCTACCA...GGTGTATGTGGTGTGC 110674

93 TyrGlyIlyTyrGlyIlyTyrGlyIlyTyrFAspGlnGlyPheGlySerAl 109

110673 TACTCAGCTATATACCGCTACGAGGC.....AGTCAATTCGAGCTGC 110630

109 aTyr.....GlyGlyTyrProGlyTyr.....T 117

110629 CTACTATTCACAGAGAGGATTACTCGCTTGTGGACACAGCGCATCAC 110580

117 yGlyIlyTyrTyrProSerGly.....TyrGlyGlyGly 128

110579 ACGAGAGATCATTCGCGCGGTGATCGTACCAATCATGAGCGGATATCT 110530

129 TyrGlyIlySerTyr.....G 134

110529 TATGGCGGCACATATGTCAGTCACTGACTCGATATTCAGCGAG 110480

134 yGlySerTyrGlyIly 139

110479 TGGTGGCTATGAGGC 110464

seq_name: gb_in:AC009742

seq_documentation_block:

LOCUS AC009742 178100 bp DNA INV 14-MAR-2001

DEFINITION Drosophila melanogaster, chromosome 3R, region 90C-90C, BAC clone BACR21K03, complete sequence.

ACCESSION AC009742

VERSION AC009742.4 GI:13324758

KEYWORDS HTG.

SOURCE fruit fly

ORGANISM Drosophila melanogaster

REFERENCE Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

AUTHORS 1 (bases 1 to 178100)

Celniker, S.E., Adams, M.D., Krommler, B., Tyler, D., Wan, K.H., Holt, R.A., Evans, C.A., Gocayne, J.D., Amanatides, P.G., Brandon, R.C., Rogers, Y., An, H., Baldwin, D., Bazon, J., Beeson, K.Y., Busam, D.A., Carlson, J.W., Center, A., Champs, M., Davenport, L.B., Dietz, S.M., Dodson, K., Dorsett, V., Doup, L.E., Doyle, C., Dresnek, D., Farfan, D., Ferreira, S., Frisoe, E., Galle, R.F., Garg, N.S., George, R.A., Gonzalez, M., Houck, J., Hoskins, R.A., Hostin, D., Howland, T.J., Ibegwam, C., Jalali, M., Kruse, D., Li, P., Mattei, B., Moshrefi, A., McIntosh, T.C., Moy, M., Murphy, B., Nelson, C., Nelson, K.A., Nunoo, J., Pacleb, J., Paragas, V., Park, S., Patel, S., Pfeiffer, B., Phouanavong, S., Pittman, G.S., Puri, V., Richards, S., Scheeler, F., Stapleton, M., Strong, R., Svirskaas, R., Tector, C., Williams, S.M.,

Zaveri,J.S., Zhan,M., Zhang,G., Zhao,Q., Zheng,L., Zheng,X.H.,
Zhong,F.N., Zhong,W., Zhou,X., Zhu,S., Zhu,X., Smith,H.O.,
Gibbs,R.A., Myers,E.W., Rubin,G.M. and Venter,J.C.
The genome sequence of *Drosophila melanogaster*
JOURNAL
MEDLINE
20196006
2 (bases 1 to 253176)
REFERENCE
Adams,M.D., Celniker,S.E., Gibbs,R.A., Rubin,G.M. and Venter,C.J.
Direct Submission
JOURNAL
Submitted (21-Mar-2000) Celera Genomics, 45 West Gude Drive,
Rockville, MD, USA
On Oct 9, 2000 this sequence version replaced gi:7300262.
COMMENT
FEATURES
source
1. 253176
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/chromosome="3R"
complement(join(<9015..9066,9230..9493,9803..10067,
10198..10425,10609..10857,11045..11208,11695..12010,
12723..12848,13260..13634,13863..14042,14361..14516,
14593..14901,15656..15799,15889..16144,16308..16518,
16861..17087,17202..17380,17568..17597,17972..18091,
18411..18546,18901..19046,19141..19290,19363..19494,
21408..21612,21940..22040,22140..22267,23854..24077,
24436..24553,25877..25997,26180..26329,26586..26747,
27102..27144))
/gene="CG12536"
/note="CG12536 gene product"
/codon_start=1
/db_xref="FLYBASE:FBan0012536"
/db_xref="FLYBASE:FBgn0038513"
/protein_id="AAF5425.1"
/db_xref="GI:7300263"
/translation="MRSRRRSGTAVLRQPHVHENTEVYLGNAIKCAIPEY
VRPYRVASWIRGEITLPLSDVAGRVVYLAASGLYRSVRSEDEGLMKISCLVTNT
LNGERKRDVAWLOVELKSLNLAPEPTTKPYMEHVERGNVHLPCNTQGNPEFLTW
YRVSALALYPIPSORVYLSRTLLIKNADBRDAGKWIQASNGFORIRLISVN
SYSVHILPQVQIVNSGCTANCTTGSIDAIDILHNGKPLQANNAITLGRDNIR
LSKSLVONQRRDGRGVYQCLVENORASQAOMAEIKLQDTPVELITYFIIONRPGP
LISLKCSASGSPPOFAWLLDSOPTMDVSLHHRALIQGFVMSGDVLSLMSIYVRPD
DGLKLCYASMSKGSVOHARLANYGPYVRAIGPIKAVASGEDITVGPFGYVEOI
RMEKHHQELTTSNHELASVADGOLVKNVPGRDOGITVCIVSRSGEARDMOL
NVNPSVLEPFRPKNLQEGRAQITTCVSSGDMPIYFSMKDDSISSJQIIEKME
EFYSLVFKDISARHSGKYCYASNAAKVNTAEIQVRAPRMRYEMPDAMIAGNT
ISINCEAGYPIPTTWKGGGSKDKFKPLKPAFEEVARNISRRNDPTLCHAK
GDEPTTIGMTONGRIDLNFRPSTIAEMKTEKGERPDPSHLETFEVSRRVYKSWR
PFDSNSPVLSLVOYQAKTYIQSHGSLAAAGDMGHVINSLESTISRSYSDSLRE
SAIVAGLTPTATFLIRMOALNIEERSATVEALVLTQEAPEARSNVOYTGGESEL
IVTQYIIPRESMNGELIGYTVNCESEKONIFISVNSLSKTIISVGWATTKATIRGL
RKYSRYATPIRAMNSFGSPMSAAIFGTAGVEAPONVNCALSSQSLKATIRGL
PLOGHGITIGYKILYRPIVHVIDEPAKLETSNLETYHLTHKASNSYIRVLAAY
ATGGGLASHPLFCOTDDVPPDAPAIKAAATLADSILISMLTPKRNCTIISHYYSR
EAGRGQAKTHVRVDENGYVTFESRSLANOMEFVNSVSTSVGESEPTSVLAQAT
NTRAPARIASFGOVRKAVGTGLVLECLAVNENPTPRARMLTRDRVTFSPYEATNEG
NLKIHREGSLSGNYTANNLFGSDEIOYVIAMKPPSAPOIIVQYASASIRVMSD

APDGAFLQCYTTSYHTAGESWSTITELLPPENMAFTISGLCGNOYTIKMSAHNVGS
GVASEEIIWTVTKGASQAPNANDELATVATVNLKLSQMWGCGSIHFFSIEHPLIGD
IRMYTVSDIDISNAEENRNLIECDPLPAQVQLRISATNDAGKTEHGFSTPHDGI
TIPSPVSPSPENDMLNNLINSTNPSGWMPTLIVYVITYSIITATITLTHRRITCG
PIABGESRTLPDGYKEDHENRNOQVYSAAPVKTVDKGNSEMEIPIYATFESYNGC
RTGAPAKTPRAVAQAQPLDTWQKFTGHEGELNNTAVPLLPSSSGFHVKSXSSW
HKORYNNTEDSTLSKSMITIVAGSQAGSKKNSNGRSKSSASGSAVAGSEDPDIS
PSTFESNMPYRVPCKSSRSSDGRAVYMPDPDSSTESNDQSPAPERRHNTPHVL
GGMAMGLGGGGGGGGGNGPAERKRSAGORSKRHGSAAQOPNOTLERRKPCGSNS
LDSVDATETASIASIAAMSGAGADLSCGFPRPAFGHGREPDGSDDERERAL
DLEQVWVESTGDQLAKMDEBELTSLARHEKKEQRELYTTHV"
complement(join(31465..31726,32418..32700,38192..38335,
39745..39767))
/gene="CG8106"
/product="CT24226"
/db_xref="FLYBASE:FBan0008106"
/db_xref="FLYBASE:FBgn0038514"
complement(<31465..>39767)
/gene="CG8106"
/map="90B1-90B3"
/db_xref="FLYBASE:FBan0008106"
/db_xref="FLYBASE:FBgn0038514"
complement(join(31503..31726,32418..32457))
/gene="CG8106"
/note="CG8106 gene product"
/codon_start=1
/db_xref="FLYBASE:FBan0008106"
/db_xref="FLYBASE:FBgn0038514"
/protein_id="AAF55426.1"
/db_xref="GI:7300264"
/translation="MSLPEPCVLVLEKALLGALLAVLATPPLATSGLRVPTLLEP
APRLFLGNDTGAQVCTVHAGSPPLVTVLVLDGSLATVPELR"
complement(join(66536..67410,67471..67541,67834..68268,
68330..68753))
/gene="CG5823"
/product="CT18279"
/db_xref="FLYBASE:FBan0005823"
/db_xref="FLYBASE:FBgn0038515"
complement(<66536..>68753)
/gene="CG5823"
/map="90B3-90B4"
/db_xref="FLYBASE:FBan0005823"
/db_xref="FLYBASE:FBgn0038515"
complement(join(66931..67410,67471..67541,67834..68134))
/gene="CG5823"
/note="CG5823 gene product"
/codon_start=1
/db_xref="FLYBASE:FBan0005823"
/db_xref="FLYBASE:FBgn0038515"
/protein_id="AAF55427.2"
/db_xref="GI:10726583"
/translation="MSSSTSGGKKOPPAVRSMKODYRLKRDPLPYTAPLENNIL
EMHVCVGPEDSPYGGYTHGTLPLPRFEPKPSITVLTNGRKRTVRLCLISDR
HBDVWNPWCVGTILTGILSLPMLSTPLTGISSENDDKQMFQAKSLFNLRNIFCE
LPEIPEIIEIKORLQTOAAVVDPAAGVRSRGLANGASIASDGNLAVAGLANPL
NAADAIDGGAAGQADLASGGAAYQNSARNSYLMWQSVYSLVITICPAIPALIVNY
IKLNQOE"
join(69939..70062,70244..70909,70982..>71058)
/gene="CG5840"
/product="CT18319"
/db_xref="FLYBASE:FBan0005840"
/db_xref="FLYBASE:FBgn0038516"
<69939..>71058
/gene="CG5840"
/map="90B3-90B4"
/db_xref="FLYBASE:FBan0005840"
/db_xref="FLYBASE:FBgn0038516"
join(69984..70062,70244..70909,70982..71058)
/gene="CG5840"
/note="CG5840 gene product"
/codon_start=1
/db_xref="FLYBASE:FBan0005840"
/db_xref="FLYBASE:FBgn0038516"


```

136 rTyrGlyGlySerTyr 141
      |||
      356 ATATGGCTCTGCCTTC 371
seq_name: gb_hcg:AP003507

seq_documentation_block:
LOCUS      AP003507      149040 bp      DNA      HTG      12-APR-2001
DEFINITION Oryza sativa chromosome 6 clone P0468A12, *** SEQUENCING IN
PROGRESS ***, in ordered pieces.
ACCESSION  AP003507
VERSION     AP003507.1  GI:13603466
KEYWORDS   HTG; HTGS_PHASE2.
SOURCE      Oryza sativa (cultivar:Nipponbare) DNA, clone:P0468A12.
ORGANISM    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
              Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE   1 (sites)
AUTHORS     Sasaki,T., Matsumoto,T. and Yamamoto,K.
TITLE       Oryza sativa nipponbare(GA3) genomic DNA, chromosome 6, PAC
            clone:P0468A12
JOURNAL     Published Only in Database (2001) In press
REFERENCE   2 (bases 1 to 149040)
AUTHORS     Sasaki,T., Matsumoto,T. and Yamamoto,K.
TITLE       Direct Submission
JOURNAL     Submitted (11-APR-2001) Takuji Sasaki, National Institute of
            Agricultural Resources, Rice Genome Research Program, Kannondai
            2-1-2, Tsukuba, Ibaraki 305-8602, Japan
            (E-mail:tsasaki@affrc.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,
            Tel:81-298-38-7441, Fax:81-298-38-7468)
NOTE        It currently consists of 1 contigs. Gaps between the contigs
            are represented as runs of N. The order of the pieces is believed
            to be correct as given, however the sizes of the gaps between them
            are based on estimates that have provided by the submitter. This
            sequence will be replaced by the finished sequence as soon as it is
            available and the accession number will be preserved.
            * NOTE: This is a 'working draft' sequence.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.
FEATURES
    source
        1..149040
           /organism="Oryza sativa"
           /cultivar="Nipponbare"
           /db_xref="taxon:4530"
           /chromosome="6"
           /clone="P0468A12"
BASE COUNT  41830 a 33432 c 32383 g 40945 t 450 others
ORIGIN
alignment_scores:
    Quality: 212.00      Length: 103
    Ratio: 3.212         Gaps: 5
    Percent Similarity: 64.078      Percent Identity: 49.515
alignment_block:
US-09-554-547-16 x AP003507 ..
Align seg 1/1 to: AP003507 from: 1 to: 149040
47 ProArGlyTYrProGlyLeuThrAlaSerIleGlyGluValGlyAl 63
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
109383 CCTACGGAGGCGCCAGCGGTTGAGATCAAAAGTGGGAGGT...GGATA 109429
   63 aArGleuGlyArgAlaGlyValGlyValSerSerTyrGly.....T 78
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
109430 CTAACCTGGTGGAGGATATGATATGT...GGTGGTACGGTGGAGGANT 109476
   78 yTyrGlyTyrProSerTyrGlyTyrProTyrGlyGlyTyrGlyTyrGly 94
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

109477 AAGTCCTCCTGGGTACGGTGTGGGTACGGTGCAGGATATGCTATCCT 109526
      |||
      95 GlyTyrGlyGly.....TyrGlyGlyTyrAspGlnGlyPheGlySerA 109
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
109527 GCGTATGCTGTGGTGCATACGGCGGTGATATGACGTGGGATGCTGCCG 109576
      |||
      109 aTyrGlyGlyTyrProGlyTyrTyrTyrTyrTyrTyrProSerGlyTyrG 126
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
109577 GATACGGTGTGGTGGCGGTGAGATATGCGGTGATACGAGCGGTGATACG 109626
      |||
      126 yGlyGlyTyr.....GlyGlySerTyrGlyGlySerTyrGlyGlySer 140
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
109627 GAGGTGATACGGTGTGGTGGGTTACGGTGCAGGATATGTTGTGGT 109676
      |||
      141 TyrThrTyr 143
      |||
109677 GGTGGTAC 109685
seq_name: gb_in:DROEDG91B

seq_documentation_block:
LOCUS      DROEDG91B      592 bp      mRNA      INV      26-APR-1993
DEFINITION Drosophila melanogaster EDG-1 cuticle protein mRNA, complete cds.
ACCESSION  M71251
VERSION     M71251.1  GI:157328
KEYWORDS   cuticle protein.
SOURCE      Drosophila melanogaster cDNA to mRNA.
ORGANISM    Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
              Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
              Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1 (bases 1 to 592)
AUTHORS     Apple,R.T. and Fristrom,J.W.
TITLE       20-hydroxyecdysone is required for, and negatively regulates,
            transcription of Drosophila pupal cuticle protein genes
            Dev. Biol. 146, 569-582 (1991)
JOURNAL     MEDLINE 91323677
FEATURES
    source
        1..592
           /organism="Drosophila melanogaster"
           /db_xref="taxon:7227"
           /gene="EDG-91"
           /gene="EDG-91"
           /gene="EDG-91"
           /note="cuticle protein"
           /codon_start=1
           /protein_id="AAA28503.1"
           /db_xref="GI:157329"
           /translation="MALVRRSCMLALLLIAGQGAAEVKTEGRTLLGLGGFGGSGVL
            SAGIGVGGLYSGFGGCGYPGGYASGYPGGYSGYNGYSGSGFGGYYPGGYYSG
            FGHRRHYHGGYPPGCGSYHNQCGSYGSHYSQSYSNQYGGGYYGGGCGNPFGR"
            577..583
polyA_signal
BASE COUNT  126 a 120 c 198 g 148 t
ORIGIN
alignment_scores:
    Quality: 210.50      Length: 138
    Ratio: 2.699         Gaps: 10
    Percent Similarity: 56.522      Percent Identity: 42.754
alignment_block:
US-09-554-547-16 x DROEDG91B ..
Align seg 1/1 to: DROEDG91B from: 1 to: 592
38 LeuTyrAlaGlyYLleGluArgAlaGlyProArgGlyTyrProGlyLeuTh 54
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
74 CTGATTGCCGCTCAGAGTCAGGCGCGGTGAAGACCAAGGTCGCAC 123
   54 rAlaSer...IleGlyGlyGluValGlyAlaArgLeuGlyArgAlaG 70

```


124 CTTGGCCCTTCTGGCGGTGATTTGTGGCAGTACGACTTACTGCCG 173
70 LVVALGlyValSerSer.....TyrGlyTyrGlyTyr 80
174 GCATCGAGTGGGTGCTGGCTGTATACGGCTTTCGAGGCGCTGCTAT 223
81 Pro.....SerTrpGlyTyrProTyrGlyGlyTyr...GlyGlyTyr 93
224 CCGTGGCTATGCGAGTGGCTACCCA...GGTGGATATGGTGGTGGCTA 270
93 TGLYTYRGLYTYRGLYTYRGLYTYRASPGLNGLYPHEGLYSERATAT 110
271 CTCACGCTATACGGCTACGAGGCG.....AGTGATTCGAGGTGGCT 314
110 YF.....GlyGlyTyr.....ProGlyTyrTyr 117
315 ACTATTCAGAGAGAGAGTTCGCGCTTTCGACACAGCGCCGATTACCCAC 364
118 GLYTYRTRYTYRPROSERGLY.....TYRGLYGLYGLYTY 129
365 GAGGATATCTATCCGGCGGCTGATCGTACACATCAGGCGGCTTCTTA 414
129 TGLYTYRGLYTYR.....GLYGLY 135
415 TCGCGCGCCACTATAGTACAGTACAGTACGAATGATATTACGAGGTG 464
135 YSerTyrGlyGly 139
465 GTCCTATCGAGCG 478
seq_name: gb_p1:AB009052

seq_documentation_block:
LOCUS AB009052 85992 bp DNA PLN 27-DEC-2000
DEFINITION Arabidopsis thaliana genomic DNA, chromosome 5, p1 clone:MNF13.
ACCESSION AB009052 BA000015
VERSION AB009052.1 GI:2656028
KEYWORDS
SOURCE

ORGANISM Arabidopsis thaliana (strain:Columbia) DNA, clone_lib:Mitsui P1
clone:MNF13.
Arabidopsis thaliana
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (sites)
Sato, S., Kaneko, T., Kotani, H., Nakamura, Y., Asamizu, E., Miyajima, N.
and Tabata, S.
Structural analysis of Arabidopsis thaliana chromosome 5. IV.
Sequence features of the regions of 1,456,315 bp covered by
nineteen physically assigned P1 and TAC clones
JOURNAL DNA Res. 5 (1), 41-54 (1998)
MEDLINE 98290546
REFERENCE 2 (bases 1 to 85992)
AUTHORS Nakamura, Y.
TITLE Direct Submission
INSTITUTION Submitted (27-NOV-1997) Yasukazu Nakamura, Kazusa DNA Research
Institute, Department of Plant Gene Research, 1532-3, Yana,
Kisarazu, Chiba 292-0612, Japan (E-mail: ynakamu@kazusa.or.jp,
Tel:81-438-52-3935, Fax:81-438-52-3934)
ADDRESS For correspondence: kaos@kazusa.or.jp
http://www.kazusa.or.jp/kaos/cgi-bin/sgd/graph.cgi?c=MNF13
For the latest information on annotation of this clone, please see
http://www.kazusa.or.jp/kaos/cgi-bin/sgd/graph.cgi?c=MNF13
Genes with similarity to proteins in the databases are described in
'product' or 'note' qualifiers. Genes that have no significant
protein similarity are described as 'unknown protein'.
The software programs used to predict genes include: Grail
(Informatics Group, Oak Ridge National Laboratory,
http://compo.ornl.gov/Grail-1.3/),
GENSCAN (Chris Burge, MIT, http://CCR-081.mit.edu/GENSCAN.html),
NetGene2 (S.M. Hebsgaard, et al., CBS, Technical University of
Denmark, http://www.cbs.dtu.dk/services/NetGene2/) and
SplicePredictor (Volker Brendel, Stanford University,

FEATURES
source

http://gremlini.zool.jastate.edu/cgi-bin/sp.cgi).
Genes encoding tRNAs are predicted by tRNAscan-SE
(Sean Eddy, Washington University School of Medicine, St. Louis,
http://genome.wustl.edu/eddy/tRNAscan-SE/).
This sequence may not be the entire insert of this clone. It may be
shorter because we remove overlaps between neighboring submissions.
The 5' clone is K21116 and the 3' clone is K1816.
location/Qualifiers

CDS

1..85992
/organism="Arabidopsis thaliana"
/strain="Columbia"
/db_xref="taxon:3702"
/chromosome="5"
/clone="MNF13"
/clone_lib="Mitsui P1"
join(244..368,668..708,812..975,1066..1165,1255..1424,
1513..2184)
/note="contains similarity to RNA binding protein
gene_id:MNF13.1"
/codon_start=1
/evidence=not experimental
/protein_id="BAB08520.1"
/db_xref="GI:9758076"
/translation="MDPYGTETVIEDEIHDPKRSIEDDDDKSQPHSGGVDSAGKI
FVGLAETTSAEFLKHFGRYGETTDSVIMKDRKQPRGFGVTYADSSVNDKVID
NHIIIGKQVEIKRTIPRGSMSNDVFKKIFVGIPSSVDDEKFEFMFGELKEHO
IMRDSHGRSGRGFGVYSEMDVDFILAKGNRLTSLGTOVELIKKAPKPNVTPPS
KRFGDSRSNFGCGYGDGCGGCGGCGYRPGPFRKSGGTYGSGSGTYGSGFGGCG
GGYGGGVGVPYREPALVSGRYGGGGGYNRKGYSNGGGGTYGGGGGDMYVGSYGEP
GGYGGPSSYGGYGGSSIGYGGYGGGGGAGGGYRGGGGYDMGGGGVGGGGYAGAG
GNGGGSFYGGGGRGGYGGGSGRYHPYR"
join(4110..4271,4372..4445,4682..4987,5065..5148,
5256..5559)
/note="gene_id:MNF13.2"
/codon_start=1
/evidence=not experimental
/product="sucrose cleavage protein-like"
/protein_id="BAB08521.1"
/db_xref="GI:9758077"
/translation="MYSTNITANSTTSARAVFLYKTPKPMALSHVEEGLPORFATIL
KDRSDLVQTKLVNCEGGSGDGVLLFPDMIRKGVKDDVDEGFEDVLYNKGPMSS
GIQBEISGTFVFCVTHASRDKRCGVCGLVILERRKQIEGRGLSDQITTKRSHVGH
KYAGNLIIFCPDSAGKITGMWGYVPDDVPELIDQIAGETIORTWRQMLPGGE
AEVYHEKVIIPNGGVYKESKGTGCGGSGNSGCCQDETFKPEPIKREVKCTIIV
FQPLDKRELYIGAIVGATATAMAFPFRRSG"
complement(join(6276..6315,6416..6509,6717..8472,
8530..8583,8670..8710))
/note="gene_id:MNF13.3
unknown protein"
/codon_start=1
/evidence=not experimental
/protein_id="BAB08522.1"
/db_xref="GI:9758078"
/translation="MSDSTKDHPPGKTEFPESILREDIDFETIKSRVEAFECNK
QGVNVQOQKELKFCDELEVRKIDSJTNKGLVYLMIILTGGSSTPEKTLKMKIEIR
DSVRDEGKKEGKEIINOHVSDANREDCRMILGRTPLSHSRDAKVL
NELDSTORFLRAMKRLKSGRMIPOLKSPRPPSPRTKLQMEIYKAVLLKSVGRN
KLOBQALAKASVDDSLKSPYKTAAPTPPSPSPRTKLQMEIYKAVLLKSVGRN
ELRHLICLDPPEAVSDLSRSVRKLTLYLBCSDLDTIPSLMEALSIVSRGN
VEHKVCREAIEETECILNVAQVKIFQCIPNELDDFDADAMEDLESDNDND
DDDDDDDDENCNLFDEKLGKNEESCRNIKLEVDOTDAMESSDSHESGACGL
VLPTDSTHETNDFESSYNRVVRPLPSITRVYPRSLVPTSMKSTIVISRDH
AMKTRKVEKRDIMEVDNDFSSRSLSVEENIKSDHGEQKPRRNKNVLAIOETSD
ETSLVANHILGRLEKRDQORNLNLEPDESILGSESRDQEKVESEKQASSQASD
EATLSVYKQITSLSESVLRLKLEMS"
join(9312..9512,9598..9659,9739..9790,9876..9990,
10076..10170,10260..10366,10460..10688,10790..10861)
/note="contains similarity to unknown protein
dbj|BA23705.1"
gene_id:MNF13.4"
/codon_start=1
/evidence=not experimental
/protein_id="BAB08523.1"

CDS

dbj|BA23705.1
gene_id:MNF13.4"
/codon_start=1
/evidence=not experimental
/protein_id="BAB08523.1"

[illegible]

23922..24026,24126..24195,24277..24373,24499..24640,
23069..25212)/
/note="gb|AAC32067.1
gene_id:MNF13.10"
/codon_start=1
/evidence=not experimental
/product="20S proteasome beta subunit PBB2"
/protein_id="BAB08528.1"
/db_xref="GI:9758084"
/translation="MSQSSVDIPPKGFSFDLCKRNDMLQKGLAPSPFKTGTIVG
LIIFDGVILGADPTATTEPIVADNCKEIHVMNPNIACCAGTAADAEAYIDWVSOL
RLHFGYGRDSRVATLTLLKKHLFSGVGHSAVALYGVDIGDPHLHTIYPGSNT
LPRTMSSGSLAASVBEAKTKBSUTDEBRIKVAEKICGIRPDLGSSNVDCVT
KGHEYLRTMEBPMPRTYSSKTSFKEIKTEVLTLITPLREIVEAGEAMEE-
complement(26370..27074)

CDS

alignment_scores:
Quality: 207.50 Length: 148
Ratio: 2.530 Gaps: 7
Percent Similarity: 55.405 Percent Identity: 40.541

alignment_block:
US-09-554-547-16 x AB009052 ..

Align seg 1/1 to: AB009052 from: 1 to: 85992

23 ArgLeuGlySerAspLeuAspThrPheGlyArgValHisGlyAsnLeuTy 39
|||||||::: ::::| ||| |
1570 CGTGCGGTACTCTCCCTACCTTTGGTGGAGATTAGAGANGTTA 1619
39 rAlaGlyIleGluArgAlaGlyProArgGlyTyrrProGlyLeuThrAlas 56
|:::||| ::::| ||| |||
1620 TGGTGC GGACACGCGTAGGATATGTCGTCAGGTGCTCCTTACAAGT 1669
56 erIleGlyGlyIleValGlyAlaArgLeuGlyGlyArgAlaGlyValGly 72
||| ||||| |:::| |:::| |:::| |
1670 CAGGTGGCGGTATAGGAGGTGGCGGTCTGTGGTTATGAGGAGTACGA 1719
73 ValSerSerTyrrGlyTyrrGlyTyrrProSerTPGlyTyrrProTyrrGlyGl 89
::: |:::| |:::| |:::| |
1720 GGAGAAATTCGGTGGCATGCCGCGTGGTGCATATGCT.....GGTGC 1760
89 yTyrrGlyGlyTyrr..... 93
||| |||
1761 TGTGGACCTTATTAGAGCAACACAGCTCTGCGATATCTGTCGTTATG 1810
94 .. GlyGlyTyrrGlyGlyTyrr..... GlyGlyTyrrAspGlnIleGly 107
||||| ||||| |:::| |:::| |
1811 GAGGAGAGCGCGGTGTTACAATAAGAGTGAAATAGATGGAGGGGCGC 1860
108 SerAlaTyrrGlyGlyTyrrProGly...TyrrTyrrGlyTyrrTyrr..... 121
:::::| ||||| |:::| |
1861 GGTGGGTACGGTGGTGGACCAAGTGTATGTGTGGTGGTCTACAGGTGA 1910
122 ProSerGly...TyrrGlyGlyGlyTyrrG 130
||| ||| |:::| |
1911 ACCAGAGAGTGTATTAGTGGACCTAATGACGATTATGAGGCGGGATTG 1960
130]y..... GlySerTyrrGlyGlySerTyrrGlyGlySer 140
||| |:::| |:::| |
1961 GTAGTAGCGGAATAGGTGGTATGTTGGTGTGAGTAATGGTGTGCT 2004

seq_name: gb_hlg:AC017979

seq_documentation_block:
LOCUS AC017979 23547 bp DNA HMG 09-DEC-1999
DEFINITION Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***, In ordered
pieces.
ACCESSION AC017979
VERSION AC017979.1 GI:6553211
KEYWORDS HTG; HTGS_PHASE2.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster

```

REFERENCE      Euarctota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
AUTHORS        Adams, M. and Venter, J.C.
TITLE          Direct Submission
JOURNAL        Submitted (09-DEC-1999) Celera Genomics, 45 West Gude Drive, Rockville, MD, USA
COMMENT         This sequence was identified as CDM:10212867 by the submitter. For more information on this record e-mail to fly@celera.com.
                * NOTE: This is a "working draft" sequence.
                * This sequence will be replaced
                  by the finished sequence as soon as it is available and
                  * the accession number will be preserved.
FEATURES       Location/Qualifiers
               source            1..23547
                                   /organism="Drosophila melanogaster"
                                   /db_xref="taxon:7227"
BASE COUNT     6844 a    4896 c   4767 g   7040 t
ORIGIN
Alignment_scores:
              Quality: 199.50           Length: 92
                 Ratio: 3.440             Gaps: 6
Percent Similarity: 63.043 Percent Identity: 50.000
alignment_block:
US-09-554-547-16 x AC017979/rev ..
$ Align seg 1/1 to reverse of: AC017979 from: 1 to: 23547

    69 AAGGlyVal...GLyAlserSerTYrGlYTyRglTYrProSerTrpgI 84
       ||| | | | : : : : : : : : : : : : : : : : : : : : 
18369 GC GGGA GTGTGGATGCCCCCACTTGGATTGATGCACCA..... 18328
    84 YTYPrOTyrGlgLYlTYrGlgLygLYrgLylTYrGYlgLY..... 98
       ||||| | | | | | | | | | | | | | | | | | | | | |
18327 ....CCATATGCGAGATACGGAGCATACGGAGCATTATGAGAAGACTATG 18282
    99 ..... TYrGlgLYtyrAspgInglyPheGly 107
       ||||| | | | | | | | | | | | | | | | | | | | | |
18281 CC TCCG CCTTC TC CGCACG AG CAT GCCGGAGAGTA TT .. CAAG AT TGCGA 18235
    108 SerAlaTYrGlgLYlTYrPGlYTyrTYrGlgLYtYrTYrTProSERGI 124
       || | | | | | | | | | | | | | | | | | | | | | |
18234 .....TACCAGCATACGGGGGATATGAGAGATA C.... .CGAGG 18200
    124 YTyrGlgLYglYTyrGlgLYseTYrGlgLYseTYrGlgLYseT 141
       |: : : : | | | | | | | | | | | | | | | | | | | |
18199 ATT TGGTGCGAGATACGGTGAGCGCTATGCGACCATTTACAACCACTTTA 18150
    141 yTYrThTYrProAnvalArGaLaSeR 149
       ::::| | | | : : : | |
18149 GCAACTAACAACTACGGAATCTTCT 18124

seq_name: gb_1n:AC009461

seq_documentation_block:
LOCUS       AC009461 171612 bp DNA INV 23-MAR-2001
DEFINITION Drosophila melanogaster, chromosome 3R, region 90B-90C, BAC clone
BACRI1113, complete sequence.
ACCESSION  AC009461
VERSION    AC009461.5 GI:13435227
KEYWORDS   HTG.
SOURCE     fruit fly.
ORGANISM   Drosophila melanogaster
Eukaryota; Metazoa; Artthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 171612)
Celisner,S.E., Adams,M.D., Kronmiller,B., Tyler,D., Wan,K.H.,
Holt,R.A., Evans,C.A., Goessyne,J.D., Aminatides,P.G., Brandon,R.C.
```

TITLE
JOURNAL
REFERENCE
AUTHORS

Unpublished
2 (bases 1 to 171612)

Cealiker,S.E., Agbayan,A., Arcaina,T.T., Baxter,E., Blazet,R.G.,
Butenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,U.,
Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L.,
Hoskins,R.A., Houston,K.A., Humstad,S.R., Katta,K., Kearney,L.,
Kim,E., Lee,B., Lewis,S., Li,P., Lomoton,M.A., Mazda,P.,
Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S.,
Pfeiffer,B., Poon,L., Sequeira,A., Seth,H., Snir,E.,
Svirskaas,R.R., Wan,K.H., Weinburg,T., Zhang,R., Zieran,L.L. and
Rubin,G.M.

DIRECT SUBMISSION
Submitted (23-AUG-1999) Drosophila genome center, Lawrence Berkeley
Laboratory, MS 64-121, Berkeley, CA 94720, USA
On Mar 23, 2001 this sequence version replaced gi:5836088.
Sequence submitted by:
Berkeley Drosophila Genome Project
Lawrence Berkeley National Laboratory, MS 64-121
Berkeley, CA 94720

This sequence was assembled using end sequences from a whole genome
shotgun and from subclones of this BAC and its neighboring clones.
For further information about this sequence, including its location
and relationship to other sequences, please visit our sequence
archive Web site (<http://www.fruitfly.org/sequence/>) or send email
to bdb@fruitfly.berkeley.edu.

location/Qualifiers
 1..171612
 /oranism="Drosophila melanogaster"
 /strain="y: cn bw sp"
 /db_xref="taxon:7227"
 /chromosome="3R"
 /map="90B-90C"
 /cclone="BACR11L13 (D981)"
 /clone_lib="RP11-98 (Rosewell) Park Cancer Institute
 Drosophila melanogaster BAC library, partial ECORI in
 PBACE3.6")

BASE COUNT 50924 a 3609 c 35092 g 49587 t

ORIGIN

alignment_scores:

Quality:	199.50	Length:	92
Ratio:	3.440	Gaps:	6
Percent Similarity:	63.043	Percent Identity:	50.000

alignment_block:
US-09-554-547-16 x AC009461 ..

Align seg 1/1 to: AC009461 from: 1 to: 171612

```

69 AlaglyVal...glyValserTyrGlyTYrGlyTYrProserTrpGl 84
|||||||      |||||      |||||      |||||      |||||
112506 GCGGAGGTGTAGTGCCGCCGAGTTGGATATGATTGCACA..... 112547

84 yTyrrPrOtyrGlyGlyTYrGlyGlyTyrrGlyGlyTYrGlyGly..... 98
|||||||      |||||      |||||      |||||      |||||
112548 ....CCATATGAGGATACGAGGATACGAGGATGTGAGGAAGCTTAG 112593

99 .....TyrglGlyTYrraspIngInglYphegIy 107
|||||||      |||||      |||||      |||||      |||||
112594 CCTCGCGCTTCGCCACGACGATGCGGGAGAGATAT...CAAGGATTTCGGA 112640

```

```

108 serAlaTyrGlyGlyTyrProGlyTyrTyrTyrTyrProserGI 124
112641 .....TACCAGGATACGGGGGATATGGAGATAC.....GGAG 112675
124 yTyrGlyGlyGlyTyrGlyGlySerTyrGlyGlySerTyrGlyGlySerT 141
112676 ATTTGGTGAGAGATACGGGTGAGGCTATGGCAGCAATACACAGTTTA 112725
141 yTThrTyrProAsnValArgAlaSer 149
112726 GCACACTACACACATACGATCTTCT 112751
seq_name: gb_in:AC009742

seq_documentation_block:
LOCUS AC009742 178100 bp DNA INV 14-MAR-2001
DEFINITION Drosophila melanogaster, chromosome 3R, region 90C-90C, BAC clone
BACR21K03, complete sequence.
ACCESSION AC009742
VERSION AC009742.4 GI:13324758
KEYWORDS HFG.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
1 (bases 1 to 178100)
Celniker,S.E., Adams,M.D., Kronmiller,B., Tyler,D., Wan,K.H.,
Holt,R.A., Evans,C.A., Gocayne,J.D., Amanalides,P.G., Brandon,R.C.,
Rogers,Y., An,H., Baldwin,D., Bazon,J., Beeson,K.Y., Busam,D.A.,
Carlson,J.W., Center,A., Champs,M., Davenport,L.B., Dietz,S.M.,
Dodson,K., Dorett,V., Doup,L.E., Doyle,C., Dresnek,D., Farfan,D.,
Fertiera,R., Fritze,E., Galle,R.F., Garg,N.S., George,R.A.,
Gonzalez,M., Houck,J., Hoskins,R.A., Hostin,D., Howland,T.J.,
Idegawa,C., Jaitai,M., Kruse,D., Li,P., Mattei,B., Moshrefi,A.,
McInosh,T.C., Moy,M., Murphy,B., Nelson,C., Nelson,K.A., Nunoo,J.,
Paciob,J., Paragas,V., Park,S., Patel,V., Pfeiffer,B.,
Phouenavong,S., Pittman,G.S., Puri,V., Richards,S., Scheeler,F.,
Stapleton,M., Strong,R., Svirskas,R., Tector,C., Williams,S.M.,
Zaveloff,T.S., Smit,H.O., Rubin,G.M. and Venter,J.C.
Sequencing of Drosophila chromosome 3R, region 90C-90C
Unpublished
2 (bases 1 to 178100)
Celniker,S.E., Aghayani,A., Arcalena,T.T., Baxter,E., Blazef,R.G.,
Butenhoff,C., Champs,M., Chaver,C., Chew,M., Ciesiolka,L.,
Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L.,
Hoskins,R.A., Houston,K.A., Hummasti,S.R., Katta,K., Kearney,L.,
Khm,E., Lee,B., Lewis,S., Li,P., Lomolan,M.A., Mazda,P.,
Moshrefi,A.R., Moshrefi,M., Nixon,K., Paclob,J.M., Park,S.,
Pfeiffer,B., Poon,L., Sequerra,A., Sethi,H., Shif,E.,
Svirskas,R.R., Wan,K.H., Weinburg,T., Zhang,R., Zieran,L.L. and
Rubin,G.M.
Direct Submission
Submitted (30-MUG-1999) Drosophila Genome Center, Lawrence Berkeley
Laboratory, MS 64-121, Berkeley, CA 94720, USA
On Mar 14, 2001 this sequence version replaced gi:5922029.
Sequence submitted by:
Berkeley Drosophila Genome Project
Lawrence Berkeley National Laboratory, MS 64-121
Berkeley, CA 94720
This sequence was assembled using end sequences from a whole genome
shotgun and from subclones of this BAC and its neighboring clones.
For further information about this sequence, including its location
and relationship to other sequences, please visit our sequence
archive Web site (http://www.fruitfly.org/sequence/) or send email
to bdg@fruitfly.berkeley.edu.
Location/Qualifiers
1. 178100
/organism="Drosophila melanogaster"
/strain="Y: cn bw sp"
/db_xref="taxon:7227"
/chromosome="3R"

FEATURES
source

```

Dunn, P., Durbin, K.J., Evangelista, C.C., Ferraz, C., Ferreira, S.,
Fleischmann, W., Foster, C., Gabriellian, A.E., Garg, N.S.,
Gelbart, W.M., Glasser, K., Glodde, A., Gong, F., Gorrell, J.H., Gu, Z.,
Guan, P., Harris, M., Harris, N.L., Harvey, D., Helman, T.J.,
Hernandez, J.R., Houck, J.J., Hostin, D., Houston, K.A., Howland, T.J.,
Jell, M.H., Idegami, C., Jaitani, M., Kalush, F., Karp, G.H., Ke, Z.,
Kennell, J.A., Ketchum, K.A., Kimmel, B.E., Kodira, C.D., Kraft, C.,
Krawitz, S., Kuip, D., Lai, Z., Laekko, P., Lei, Y., Levitsky, A.A.,
Li, J., Li, Z., Liang, Y., Lin, X., Liu, X., Maltel, B., McIntosh, T.C.,
McLeod, M.P., McPherson, D., Merkulov, G., Milshina, N.V., Mobashy, C.,
Morris, J., Moshrefi, A., Mount, S.M., Moy, M., Murphy, B., Murphy, L.,
Muzny, D.M., Nelson, D.L., Nelson, D.R., Nelson, K.A., Nixon, K.,
Nusskern, D.R., Pacle, J.W., Palazzolo, M., Peltman, G.S., Pan, S.,
Pollard, J., Puri, V., Reese, M.G., Reinert, K., Remington, K.,
Sundares, R.D., Scheeler, F., Shen, H., Shue, B.C., Siden-Kiamos, I.,
Stimpson, M., Skupski, M.P., Smith, T., Spier, E., Tector, C., Turner, R.,
Venter, E., Wang, A.H., Wang, X., Wang, Z., Wasserman, D.A.,
Weinstock, G.M., Weissbach, J., Williams, S.M., Woodage, T.,
Worley, K.C., Wu, D., Yang, S., Yao, Q.A., Ye, J., Yeh, R.F.,
Zaveri, J.S., Zhao, M., Zhao, G., Zhao, Q., Zheng, L., Zheng, X.H.,
Zhong, F.N., Zhong, W., Zhou, X., Zhu, S., Zhu, X., Smith, H.O.,
Gibbs, R.A., Myers, E.W., Rubin, G.M. and Venter, J.C.
The genome sequence of *Drosophila melanogaster*
Science 287 (5461), 2185-2195 (2000)

TITLE
JOURNAL
MEDLINE
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
SOURCE

2 (bases 1 to 253176)
Adams, M.D., Celniker, S.E., Gibbs, R.A., Rubin, G.M. and Venter, C.J.
Submitted (21-MAR-2000) Celera Genomics, 45 West Gude Drive,
Rockville, MD, USA
On Oct 9, 2000 this sequence version replaced gi:7300262.
Location/Qualifiers
1. 253176
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/chromosome="3R"
complement(join(<9015..9066,9230..9493,9803..10067,
10198..10425,10609..10857,11045..11208,11695..12010,
12723..12848,13260..13634,13863..14042,14361..14516,
14593..14901,15656..15799,15889..16144,16308..16518,
16861..17087,17202..17380,17568..17597,17972..18091,
18411..18546,18901..19046,19141..19290,19363..19494,
21408..21612,21940..22040,22140..22267,223854..24077,
24436..24553,25877..25997,26180..26329,26586..26747,
27102..>27144))
/gene="CG12536"
/product="CT33960"
/db_xref="FLYBASE:Fban0012536"
/db_xref="FLYBASE:Fbgn0038513"
complement(<9015..>27144)
/gene="CG12536"
/map="90B1-90B2"
/db_xref="FLYBASE:Fban0012536"
/db_xref="FLYBASE:Fbgn0038513"
complement(join(9015..9066,9230..9493,9803..10067,
10198..10425,10609..10857,11045..11208,11695..12010,
12723..12848,13260..13634,13863..14042,14361..14516,
14593..14901,15656..15799,15889..16144,16308..16518,
16861..17087,17202..17380,17568..17597,17972..18091,
18411..18546,18901..19046,19141..19290,19363..19494,
21408..21612,21940..22040,22140..22267,223854..24077,
24436..24553,25877..25997,26180..26329,26586..26747,
27102..>27144))
/gene="CG12536"
/note="CG12536 gene product"
/codon_start=1
/db_xref="FLYBASE:Fban0012536"
/db_xref="FLYBASE:Fbgn0038513"
/protein_id="AAFS5425.1"
/db_xref="GI:7300263"
/translation="MSLRRRAGSTAVLVROFHVHIVENTEVYLGNSALIKAIPEY
VRBYVAVSWHRGEETLLPDLSDVAGRVVLAASGDLVRSVSESDGLMFSLVNT

gene
CDS
mRNA

1. 253176
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/chromosome="3R"
complement(join(<9015..9066,9230..9493,9803..10067,
10198..10425,10609..10857,11045..11208,11695..12010,
12723..12848,13260..13634,13863..14042,14361..14516,
14593..14901,15656..15799,15889..16144,16308..16518,
16861..17087,17202..17380,17568..17597,17972..18091,
18411..18546,18901..19046,19141..19290,19363..19494,
21408..21612,21940..22040,22140..22267,223854..24077,
24436..24553,25877..25997,26180..26329,26586..26747,
27102..>27144))
/gene="CG12536"
/product="CT33960"
/db_xref="FLYBASE:Fban0012536"
/db_xref="FLYBASE:Fbgn0038513"
complement(<9015..>27144)
/gene="CG12536"
/map="90B1-90B2"
/db_xref="FLYBASE:Fban0012536"
/db_xref="FLYBASE:Fbgn0038513"
complement(join(9015..9066,9230..9493,9803..10067,
10198..10425,10609..10857,11045..11208,11695..12010,
12723..12848,13260..13634,13863..14042,14361..14516,
14593..14901,15656..15799,15889..16144,16308..16518,
16861..17087,17202..17380,17568..17597,17972..18091,
18411..18546,18901..19046,19141..19290,19363..19494,
21408..21612,21940..22040,22140..22267,223854..24077,
24436..24553,25877..25997,26180..26329,26586..26747,
27102..>27144))
/gene="CG12536"
/note="CG12536 gene product"
/codon_start=1
/db_xref="FLYBASE:Fban0012536"
/db_xref="FLYBASE:Fbgn0038513"
/protein_id="AAFS5425.1"
/db_xref="GI:7300263"
/translation="MSLRRRAGSTAVLVROFHVHIVENTEVYLGNSALIKAIPEY
VRBYVAVSWHRGEETLLPDLSDVAGRVVLAASGDLVRSVSESDGLMFSLVNT

gene
CDS
mRNA

1. 253176
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/chromosome="3R"
complement(join(<9015..9066,9230..9493,9803..10067,
10198..10425,10609..10857,11045..11208,11695..12010,
12723..12848,13260..13634,13863..14042,14361..14516,
14593..14901,15656..15799,15889..16144,16308..16518,
16861..17087,17202..17380,17568..17597,17972..18091,
18411..18546,18901..19046,19141..19290,19363..19494,
21408..21612,21940..22040,22140..22267,223854..24077,
24436..24553,25877..25997,26180..26329,26586..26747,
27102..>27144))
/gene="CG12536"
/product="CT33960"
/db_xref="FLYBASE:Fban0012536"
/db_xref="FLYBASE:Fbgn0038513"
complement(<9015..>27144)
/gene="CG12536"
/map="90B1-90B2"
/db_xref="FLYBASE:Fban0012536"
/db_xref="FLYBASE:Fbgn0038513"
complement(join(9015..9066,9230..9493,9803..10067,
10198..10425,10609..10857,11045..11208,11695..12010,
12723..12848,13260..13634,13863..14042,14361..14516,
14593..14901,15656..15799,15889..16144,16308..16518,
16861..17087,17202..17380,17568..17597,17972..18091,
18411..18546,18901..19046,19141..19290,19363..19494,
21408..21612,21940..22040,22140..22267,223854..24077,
24436..24553,25877..25997,26180..26329,26586..26747,
27102..>27144))
/gene="CG12536"
/note="CG12536 gene product"
/codon_start=1
/db_xref="FLYBASE:Fban0012536"
/db_xref="FLYBASE:Fbgn0038513"
/protein_id="AAFS5425.1"
/db_xref="GI:7300263"
/translation="MSLRRRAGSTAVLVROFHVHIVENTEVYLGNSALIKAIPEY
VRBYVAVSWHRGEETLLPDLSDVAGRVVLAASGDLVRSVSESDGLMFSLVNT

gene
CDS
mRNA

1. 253176
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/chromosome="3R"
complement(join(<9015..9066,9230..9493,9803..10067,
10198..10425,10609..10857,11045..11208,11695..12010,
12723..12848,13260..13634,13863..14042,14361..14516,
14593..14901,15656..15799,15889..16144,16308..16518,
16861..17087,17202..17380,17568..17597,17972..18091,
18411..18546,18901..19046,19141..19290,19363..19494,
21408..21612,21940..22040,22140..22267,223854..24077,
24436..24553,25877..25997,26180..26329,26586..26747,
27102..>27144))
/gene="CG12536"
/product="CT33960"
/db_xref="FLYBASE:Fban0012536"
/db_xref="FLYBASE:Fbgn0038513"
complement(<9015..>27144)
/gene="CG12536"
/map="90B1-90B2"
/db_xref="FLYBASE:Fban0012536"
/db_xref="FLYBASE:Fbgn0038513"
complement(join(9015..9066,9230..9493,9803..10067,
10198..10425,10609..10857,11045..11208,11695..12010,
12723..12848,13260..13634,13863..14042,14361..14516,
14593..14901,15656..15799,15889..16144,16308..16518,
16861..17087,17202..17380,17568..17597,17972..18091,
18411..18546,18901..19046,19141..19290,19363..19494,
21408..21612,21940..22040,22140..22267,223854..24077,
24436..24553,25877..25997,26180..26329,26586..26747,
27102..>27144))
/gene="CG12536"
/note="CG12536 gene product"
/codon_start=1
/db_xref="FLYBASE:Fban0012536"
/db_xref="FLYBASE:Fbgn0038513"
/protein_id="AAFS5425.1"
/db_xref="GI:7300263"
/translation="MSLRRRAGSTAVLVROFHVHIVENTEVYLGNSALIKAIPEY
VRBYVAVSWHRGEETLLPDLSDVAGRVVLAASGDLVRSVSESDGLMFSLVNT

gene
CDS
mRNA

1. 253176
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/chromosome="3R"
complement(join(<9015..9066,9230..9493,9803..10067,
10198..10425,10609..10857,11045..11208,11695..12010,
12723..12848,13260..13634,13863..14042,14361..14516,
14593..14901,15656..15799,15889..16144,16308..16518,
16861..17087,17202..17380,17568..17597,17972..18091,
18411..18546,18901..19046,19141..19290,19363..19494,
21408..21612,21940..22040,22140..22267,223854..24077,
24436..24553,25877..25997,26180..26329,26586..26747,
27102..>27144))
/gene="CG12536"
/product="CT33960"
/db_xref="FLYBASE:Fban0012536"
/db_xref="FLYBASE:Fbgn0038513"
complement(<9015..>27144)
/gene="CG12536"
/map="90B1-90B2"
/db_xref="FLYBASE:Fban0012536"
/db_xref="FLYBASE:Fbgn0038513"
complement(join(9015..9066,9230..9493,9803..10067,
10198..10425,10609..10857,11045..11208,11695..12010,
12723..12848,13260..13634,13863..14042,14361..14516,
14593..14901,15656..15799,15889..16144,16308..16518,
16861..17087,17202..17380,17568..17597,17972..18091,
18411..18546,18901..19046,19141..19290,19363..19494,
21408..21612,21940..22040,22140..22267,223854..24077,
24436..24553,25877..25997,26180..26329,26586..26747,
27102..>27144))
/gene="CG12536"
/note="CG12536 gene product"
/codon_start=1
/db_xref="FLYBASE:Fban0012536"
/db_xref="FLYBASE:Fbgn0038513"
/protein_id="AAFS5425.1"
/db_xref="GI:7300263"
/translation="MSLRRRAGSTAVLVROFHVHIVENTEVYLGNSALIKAIPEY
VRBYVAVSWHRGEETLLPDLSDVAGRVVLAASGDLVRSVSESDGLMFSLVNT

gene
CDS
mRNA

1. 253176
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/chromosome="3R"
complement(join(<9015..9066,9230..9493,9803..10067,
10198..10425,10609..10857,11045..11208,11695..12010,
12723..12848,13260..13634,13863..14042,14361..14516,
14593..14901,15656..15799,15889..16144,16308..16518,
16861..17087,17202..17380,17568..17597,17972..18091,
18411..18546,18901..19046,19141..19290,19363..19494,
21408..21612,21940..22040,22140..22267,223854..24077,
24436..24553,25877..25997,26180..26329,26586..26747,
27102..>27144))
/gene="CG12536"
/product="CT33960"
/db_xref="FLYBASE:Fban0012536"
/db_xref="FLYBASE:Fbgn0038513"
complement(<9015..>27144)
/gene="CG12536"
/map="90B1-90B2"
/db_xref="FLYBASE:Fban0012536"
/db_xref="FLYBASE:Fbgn0038513"
complement(join(9015..9066,9230..9493,9803..10067,
10198..10425,10609..10857,11045..11208,11695..12010,
12723..12848,13260..13634,13863..14042,14361..14516,
14593..14901,15656..15799,15889..16144,16308..16518,
16861..17087,17202..17380,17568..17597,17972..18091,
18411..18546,18901..19046,19141..19290,19363..19494,
21408..21612,21940..22040,22140..22267,223854..24077,
24436..24553,25877..25997,26180..26329,26586..26747,
27102..>27144))
/gene="CG12536"
/note="CG12536 gene product"
/codon_start=1
/db_xref="FLYBASE:Fban0012536"
/db_xref="FLYBASE:Fbgn0038513"
/protein_id="AAFS5425.1"
/db_xref="GI:7300263"
/translation="MSLRRRAGSTAVLVROFHVHIVENTEVYLGNSALIKAIPEY
VRBYVAVSWHRGEETLLPDLSDVAGRVVLAASGDLVRSVSESDGLMFSLVNT

gene
CDS
mRNA

1. 253176
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/chromosome="3R"
complement(join(<9015..9066,9230..9493,9803..10067,
10198..10425,10609..10857,11045..11208,11695..12010,
12723..12848,13260..13634,13863..14042,14361..14516,
14593..14901,15656..15799,15889..16144,16308..16518,
16861..17087,17202..17380,17568..17597,17972..18091,
18411..18546,18901..19046,19141..19290,19363..19494,
21408..21612,21940..22040,22140..22267,223854..24077,
24436..24553,25877..25997,26180..26329,26586..26747,
27102..>27144))
/gene="CG12536"
/product="CT33960"
/db_xref="FLYBASE:Fban0012536"
/db_xref="FLYBASE:Fbgn0038513"
complement(<9015..>27144)
/gene="CG12536"
/map="90B1-90B2"
/db_xref="FLYBASE:Fban0012536"
/db_xref="FLYBASE:Fbgn0038513"
complement(join(9015..9066,9230..9493,9803..10067,
10198..10425,10609..10857,11045..11208,11695..12010,
12723..12848,13260..13634,13863..14042,14361..14516,
14593..14901,15656..15799,15889..16144,16308..16518,
16861..17087,17202..17380,17568..17597,17972..18091,
18411..18546,18901..19046,19141..19290,19363..19494,
21408..21612,21940..22040,22140..22267,223854..24077,
24436..24553,25877..25997,26180..26329,26586..26747,
27102..>27144))
/gene="CG12536"
/note="CG12536 gene product"
/codon_start=1
/db_xref="FLYBASE:Fban0012536"
/db_xref="FLYBASE:Fbgn0038513"
/protein_id="AAFS5425.1"
/db_xref="GI:7300263"
/translation="MSLRRRAGSTAVLVROFHVHIVENTEVYLGNSALIKAIPEY
VRBYVAVSWHRGEETLLPDLSDVAGRVVLAASGDLVRSVSESDGLMFSLVNT

gene
CDS
mRNA

1. 253176
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/chromosome="3R"
complement(join(<9015..9066,9230..9493,9803..10067,
10198..10425,10609..10857,11045..11208,11695..12010,
12723..12848,13260..13634,13863..14042,14361..14516,
14593..14901,15656..15799,15889..16144,16308..16518,
16861..17087,17202..17380,17568..17597,17972..18091,
18411..18546,18901..19046,19141..19290,19363..19494,
21408..21612,21940..22040,22140..22267,223854..24077,
24436..24553,25877..25997,26180..26329,26586..26747,
27102..>27144))
/gene="CG12536"
/product="CT33960"
/db_xref="FLYBASE:Fban0012536"
/db_xref="FLYBASE:Fbgn0038513"
complement(<9015..>27144)
/gene="CG12536"
/map="90B1-90B2"
/db_xref="FLYBASE:Fban0012536"
/db_xref="FLYBASE:Fbgn0038513"
complement(join(9015..9066,9230..9493,9803..10067,
10198..10425,10609..10857,11045..11208,11695..12010,
12723..12848,13260..13634,13863..14042,14361..14516,
14593..14901,15656..15799,15889..16144,16308..16518,
16861..17087,17202..17380,17568..17597,17972..18091,
18411..18546,18901..19046,19141..19290,19363..19494,
21408..21612,21940..22040,22140..22267,223854..24077,
24436..24553,25877..25997,26180..26329,26586..26747,
27102..>27144))
/gene="CG12536"
/note="CG12536 gene product"
/codon_start=1
/db_xref="FLYBASE:Fban0012536"
/db_xref="FLYBASE:Fbgn0038513"
/protein_id="AAFS5425.1"
/db_xref="GI:7300263"
/translation="MSLRRRAGSTAVLVROFHVHIVENTEVYLGNSALIKAIPEY
VRBYVAVSWHRGEETLLPDLSDVAGRVVLAASGDLVRSVSESDGLMFSLVNT

gene
CDS
mRNA

1. 253176
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/chromosome="3R"
complement(join(<9015..9066,9230..9493,9803..10067,
10198..10425,10609..10857,11045..11208,11695..12010,
12723..12848,13260..13634,13863..14042,14361..14516,
14593..14901,15656..15799,15889..16144,16308..16518,
16861..17087,17202..17380,17568..17597,17972..18091,
18411..18546,18901..19046,19141..19290,19363..19494,
21408..21612,21940..22040,22140..22267,223854..24077,
24436..24553,25877..25997,26180..26329,26586..26747,
27102..>27144))
/gene="CG12536"
/product="CT33960"
/db_xref="FLYBASE:Fban0012536"
/db_xref="FLYBASE:Fbgn0038513"
complement(<9015..>27144)
/gene="CG12536"
/map="90B1-90B2"
/db_xref="FLYBASE:Fban0012536"
/db_xref="FLYBASE:Fbgn0038513"
complement(join(9015..9066,9230..9493,9803..10067,
10198..10425,10609..10857,11045..11208,11695..12010,
12723..12848,13260..13634,13863..14042,14361..14516,
14593..14901,15656..15799,15889..16144,16308..16518,
16861..17087,17202..17380,17568..17597,17972..18091,
18411..18546,18901..19046,19141..19290,19363..19494,
21408..21612,21940..22040,22140..22267,223854..24077,
24436..24553,25877..25997,26180..26329,26586..26747,
27102..>27144))
/gene="CG12536"
/note="CG12536 gene product"
/codon_start=1
/db_xref="FLYBASE:Fban0012536"
/db_xref="FLYBASE:Fbgn0038513"
/protein_id="AAFS5425.1"
/db_xref="GI:7300263"
/translation="MSLRRRAGSTAVLVROFHVHIVENTEVYLGNSALIKAIPEY
VRBYVAVSWHRGEETLLPDLSDVAGRVVLAASGDLVRSVSESDGLMFSLVNT

gene
CDS
mRNA

1. 253176
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/chromosome="3R"
complement(join(<9015..9066,9230..9493,9803..10067,
10198..10425,10609..10857,11045..11208,11695..12010,
12723..12848,13260..13634,13863..14042,14361..14516,
14593..14901,15656..15799,15889..16144,16308..16518,
16861..17087,17202..17380,17568..17597,17972..18091,
18411..18546,18901..19046,19141..19290,19363..19494,
21408..21612,21940..22040,22140..22267,223854..24077,
24436..24553,25877..25997,26180..26329,26586..26747,
27102..>27144))
/gene="CG12536"
/product="CT33960"
/db_xref="FLYBASE:Fban0012536"
/db_xref="FLYBASE:Fbgn0038513"
complement(<9015..>27144

[illegible]

OM of: US-09-554-547-16 to: N_Geneseq_1101.* out-format : pfs

Date: Feb 11, 2002 12:27 PM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-MODEL=frimex_p2n.model -DEV=xlp
-O=cgnt2.1/USPTO.spool/US09554547/rnmat_11022002.113853.17732/app_query.fasta_1.214
-DB=N_Geneseq_1101 -OPMT=fastest -SUFFIX=ring -GAPOP=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
-QGAPOP=4.500 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
-FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blsum62
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
-NORTEXT -MINLEN=0 -MAXLEN=200000000
-USER=US09554547_@CGN1_1_268 -NCPU=6 -ICPU=3 -LONGLOG -NO_XLPHY
-WAIT -THREADS=1

Search information block:

Query: US-09-554-547-16
Query length: 154
Database: N_Geneseq_1101.*
Database sequences: 930621
Database length: 428662619
Search time (sec): 146.330000

score_list:

Sequence	Strd Orig	ZScore	EScore Len	Documentation
/SID88/gcgdata/geneseq/geneseq/NA1999.DAT:AA55699			851.00	1301.35
/SID88/gcgdata/geneseq/geneseq/NA2000.DAT:AA64661			202.50	300.66
/SID88/gcgdata/geneseq/geneseq/NA2001.DAT:AA110862			180.50	279.18
/SID88/gcgdata/geneseq/geneseq/NA2001.DAT:AA132122			180.50	279.18
/SID88/gcgdata/geneseq/geneseq/NA2001.DAT:AA100786			180.50	279.18
/SID88/gcgdata/geneseq/geneseq/NA2001.DAT:AA123045			180.50	274.70
/SID88/gcgdata/geneseq/geneseq/NA2001.DAT:AA147337			180.50	274.70
/SID88/gcgdata/geneseq/geneseq/NA2001.DAT:AA107740			180.50	270.31
/SID88/gcgdata/geneseq/geneseq/NA2000.DAT:AA53029			180.50	267.91
/SID88/gcgdata/geneseq/geneseq/NA2000.DAT:AAV57370			175.50	267.91
/SID88/gcgdata/geneseq/geneseq/NA1998.DAT:AAV57370			175.50	267.91
/SID88/gcgdata/geneseq/geneseq/NA2000.DAT:AA544994			175.50	265.60
/SID88/gcgdata/geneseq/geneseq/NA2001.DAT:AA158656			175.50	262.21
/SID88/gcgdata/geneseq/geneseq/NA2001.DAT:AA189925			175.50	256.39
/SID88/gcgdata/geneseq/geneseq/NA2001.DAT:AA160442			175.50	254.94
/SID88/gcgdata/geneseq/geneseq/NA2001.DAT:AA43233			175.00	271.02
/SID88/gcgdata/geneseq/geneseq/NA2001.DAT:AAH57413			173.00	249.81
/SID88/gcgdata/geneseq/geneseq/NA2000.DAT:AA40661			173.00	245.54
/SID88/gcgdata/geneseq/geneseq/NA1995.DAT:AAQ98470			172.00	251.97
/SID88/gcgdata/geneseq/geneseq/NA2000.DAT:AA61378			171.00	266.33
/SID88/gcgdata/geneseq/geneseq/NA2000.DAT:AA61376			170.00	264.38
/SID88/gcgdata/geneseq/geneseq/NA2000.DAT:AA61377			170.00	264.38
/SID88/gcgdata/geneseq/geneseq/NA1994.DAT:AA055750			169.00	252.28
/SID88/gcgdata/geneseq/geneseq/NA2000.DAT:AA635675			167.50	258.60
/SID88/gcgdata/geneseq/geneseq/NA2000.DAT:AA632115			167.00	252.76
/SID88/gcgdata/geneseq/geneseq/NA2000.DAT:AA639441			166.50	256.10
/SID88/gcgdata/geneseq/geneseq/NA2000.DAT:AA640899			166.50	256.09
/SID88/gcgdata/geneseq/geneseq/NA2000.DAT:AA638839			166.50	249.47
/SID88/gcgdata/geneseq/geneseq/NA2000.DAT:AA651780			166.00	245.95
/SID88/gcgdata/geneseq/geneseq/NA2000.DAT:AA640433			165.00	252.35
/SID88/gcgdata/geneseq/geneseq/NA2000.DAT:AA640595			165.00	248.93
/SID88/gcgdata/geneseq/geneseq/NA2000.DAT:AA636421			164.00	252.92
/SID88/gcgdata/geneseq/geneseq/NA2000.DAT:AA642090			163.50	250.55
/SID88/gcgdata/geneseq/geneseq/NA2000.DAT:AA661380			163.00	253.93
/SID88/gcgdata/geneseq/geneseq/NA2001.DAT:AAH85637			163.00	241.71
/SID88/gcgdata/geneseq/geneseq/NA2001.DAT:AAH34552			163.00	235.91
/SID88/gcgdata/geneseq/geneseq/NA2000.DAT:AA616296			163.00	225.86
/SID88/gcgdata/geneseq/geneseq/NA2001.DAT:AA112240			162.50	251.91
/SID88/gcgdata/geneseq/geneseq/NA2001.DAT:AA133594			162.50	251.91

/SID88/gcgdata/geneseq/geneseq/NA2001.DAT:AA102153 - 162.50 251.91 5.9e-06 465
/SID88/gcgdata/geneseq/geneseq/NA1998.DAT:AAV57377 + 162.50 230.31 9.4e-05 6596
/SID88/gcgdata/geneseq/geneseq/NA2000.DAT:AA647160 + 161.50 249.94 7.6e-06 491
/SID88/gcgdata/geneseq/geneseq/NA2000.DAT:AA642445 + 161.50 247.22 1.1e-05 686
/SID88/gcgdata/geneseq/geneseq/NA2001.DAT:AA680405 + 161.50 244.44 1.5e-05 965

seq_name: /SID88/gcgdata/geneseq/geneseq/NA1999.DAT:AA55699

seq_documentation_block:

ID AAX55699 standard; CDNA; 656 BP.

AC AAX55699;

DT 30-JUL-1999 (first entry)

DE CDNA clone 64 encoding a cement protein.

XX Tissue cement protein; TCP; blood-feeding; ectoparasite; therapy;

KW pharmaceutical; vaccine; tissue bonding; surgical wound; laceration;

KW skin grafting; immunogen; infection; arthropod parasite; ss.

OS Rhinoceros appendiculatus.

XX W09924567-A1.

PN 20-MAY-1999.

XX 12-NOV-1998; 98WO-GB03397.

PR 12-NOV-1997; 97GB-0023945.

XX (NAEN) NAT ENVIRONMENTAL RES COUNCIL.

PI Nuttall Patricia A, Paesen GC;

DR WPI: 1999-327399/27.

XX P-PSDB: AAY13500.

PT Tissue cement proteins produced by blood-feeding ectoparasites and

XX related polynucleotides

XX Disclosure; Fig 7; 65pp; English.

CC The invention relates to tissue cement proteins (TCPs) produced by

CC blood-feeding ectoparasites. The TCP, in a pharmaceutical composition,

CC is useful for therapy, as a vaccine or vaccine component. The TCP itself

CC are useful for bonding animal tissues. This may be temporary or permanent

CC bonding and used in, e.g. repair of incised surgical wounds, lacerations,

CC skin grafting, etc. The TCP can also be used as a protective immunogen in

CC the control of diseases caused by infections transmitted by arthropod

CC parasites. The tick TCPs provide a non-immunogenic tissue cement capable

CC of bonding mammalian tissue with great strength. The hardening or elastic

CC properties of the TCPs can be tailored for particular requirements.

XX Sequence 656 BP; 160 A; 158 C; 191 G; 147 T; 0 other;

XX

XX

XX

XX

XX

```

|||||
97  TGCAGCAGAGGCTGCTGCTTGGAGACGACCTGGATACATTGGAGAGG 146
34  alhhsGlyAsnLeuTyraLaglylleGluArgAlaGlyProArgGlyTyr 50
|||||
147 TACACGGTATACCTATATGCGGCATCGAAGAGCTGCGCCCTCGTGATAC 196
51 ProGlyLeuThrAlaSerIleGlyGlyIleValGlyAlaArgLeuGlyI 67
|||||
197 CCAAGGCTTACCGCATCGATTGGAGGGAGTGGTGCACAGACCTGGTGG 246
67 yAArgAlaGlyValGlyValSerSerTyrglyTyrglyTyrProSerTrpG 84
|||||
247 TCGTCCGGGTGGAGTGGAGTACGACCTACGGCTATGTTACCCCTTCATGGG 296
84 lyTyrProTyrglyGlyTyrglyTyrglyTyrglyTyrglyTyrgly 100
|||||
297 GCTATCCGTATGGTGGATACGGTGGATACGGTGGATACGGTGGATACGGT 346
101 GLyTyrAspGlnGlyPheGlySerAlaTyrglyGlyTyrglyProGlyTyr 117
|||||
347 GGAATATATCAGGGTTTGGCTCTGCATACGGCGGCTACCCCGGCTACTA 396
117 rGlyTyrglyTyrProSerGlyTyrglyGlyTyrglyTyrglyTyrgly 134
|||||
397 TGGCTACTACTATCCAGTGGCTACGGTGGGCTACGGTGGTACGTAACG 446
134 lyGlySerTyrglyGlySerTyrglyTyrProAsnValArgAlaSerAla 150
|||||
447 GTGGCAGCTACGGTGGTACGTAACCTATCCCAACGTTCCGGCTTCAGCT 496
151 GLyAlaAlaAla 154
|||||
497 GGTGCCCGCAGCT 508

seq_name: /stds/gcgdata/geneseq/geneseqn/NA2000.DAT:AAA64661
seq_documentation_block:
ID AAA64661 standard; cDNA: 2126 BP.
XX
AC AAA64661;
XX
DT 02-JAN-2001 (first entry)
XX
DE cDNA encoding TBP associated factor (TAFII68).
XX
KW Melanoma; thyroid tumour; MEL3; MEL7; Thy5; Thy6; Thy11; Thy14; Thy15;
KW rectal cancer; lung cancer; breast cancer; colon cancer;
KW TBP associated factor; TAFII68; ss.
XX
OS Homo sapiens.
XX
PN WO200050595-A2.
XX
PD 31-AUG-2000.
XX
PE 25-FEB-2000; 2000WO-US04929.
XX
PR 25-FEB-1999; 99US-0257417.
XX
PA (GOUT/) GOUT I.
PA (RODN/) RODIN N.
PA (FILO/) FILOENKO V.
PA (MATS/) MATSUKA G.
PA (SCAN/) SCANLAN M.
PA (OLDL/) OLD L.
PA (BILY/) BILYNSKY B.
XX
PI Gout I, Rodin N, Filonenko V, Matsuka G, Scanlan M, Old L,
XX Bilynsky B.
XX WPI: 2000-572092/53.
DR
XX

```

```

PT Novel isolated nucleic acid molecules for diagnosing and treating
PT melanoma, thyroid tumors, rectal, lung, breast and colon cancers
XX
XX
PS Claim 17, Page 88; 94pp; English.
XX
CC The specification describes polynucleotides which are associated with
CC melanoma and thyroid tumors. Specifically, these are MEL3, MEL7, Thy5,
CC Thy6, Thy11, Thy14, and Thy15. The polynucleotides are useful for
CC diagnosing and treating a patient with melanoma, thyroid tumour,
CC rectal cancer, lung cancer, breast cancer or colon cancer. The
CC present sequence represents a polynucleotide of the invention.
XX
SQ Sequence 2126 BP; 609 A; 373 C; 691 G; 453 T; 0 other;

alignment_scores:
Quality: 202.50 Length: 138
Ratio: 2.531 Gaps: 9
Percent Similarity: 57.971 Percent Identity: 43.478

alignment_block:
US-09-554-547-16 x AAA64661 ..

Align seg 1/1 to: AAA64661 from: 1 to: 2126
20 ArgAlaGlyArgLeuGlySerAspLeuAspThrPheGlyArgValHisG1 36
|||||
1412 AGAGCGCGGCGCTATGCGGGGAC.....AGAGG 1440
36 yAsnLeuTyraLaglylleGluArgAlaGlyProArgGlyTyrProGlyL 53
|||||
1441 AGCGCGCTATGGAGGA...GACCGAGGAGGT.....GGCTATGGAGGAG 1481
53 euThrAlaSerIleGlyGlyIleValGlyAlaArgLeuGlyGly...Arg 68
|||||
1482 ATCGAGGTGGCTATGAGAGAGACCGAGGTGGAGCTATGCTATGGAGACCGA 1531
69 AlaGlyValGlyValSerSerTyrglyTyrglyTyrProSerTyrglyTy 85
|||||
1532 GGAGCTATGAGAGAGATCGAGAGGCTATGGAGAGACGCGGGCGCTATG 1581
85 r.....ProTyrglyGlyTyrglyTyrglyTyrglyTyrglyTyrg 100
|||||
1582 TGGAGGAGATCGAGAGGCTATGGAGAGACGAGAGCGGGCGCTATG 1631
100 lyGlyTyrAspGlnGlyPheGlySerAlaTyrglyGlyTyr.....Pro 114
|||||
1632 GAGGA...GACCGTGGTGGCAGTGGCTACGGTGGACACCGAAGTGA 1678
115 GLyTyrTyrglyTyrglyTyrProSerGlyTyr.....GlyG1 127
|||||
1679 GCGTATGAGAGACGAGAGTGTGGCGGCTATGGAGAGACGAGCGGTGG 1728
127 yGlyTyrgly.....GlySerTyrglyGlySerTyrglyGlySerT 141
|||||
1729 GGGCTACGAGAGACCGAGGTGCTATGGAGCAAAATGGAGGAGAA 1778
141 yThrTyrglyProAsn 145
|||||
1779 ACGACTACAGAAAT 1792

seq_name: /stds/gcgdata/geneseq/geneseqn/NA2001.DAT:AA110862
seq_documentation_block:
ID AA110862 standard; DNA: 479 BP.
XX
AC AA110862;
XX
DT 12-OCT-2001 (first entry)
XX
DE Probe #795 for gene expression analysis in human cervical cell sample.
XX Probe: human; microarray; gene expression; cervical epithelial cell;
KW

```



```
122 roserglytyrglyglytyrglyserglyserglysergly 138
      ::::::::::::::::::::
435 ATGAGCTATGCTGCTGCTGCTGCTATGCTGCTGCTGCT 484
139 Gly 139
      |||
485 GGT 487

seq_name: /SID58/gcdata/geneseq/geneseqn/NA2001.DAT:AA107740

seq_documentation_block:
ID  AA107740 standard; DNA: 830 BP.
XX
AC  AA107740;
XX
DT  09-OCT-2001 (first entry)
XX
DE  Probe #7731 used to measure gene expression in human breast sample.
XX
KW  Probe: human; breast disease; breast cancer; development disorder; ss;
XX  inflammatory disease; proliferative breast disease; non-carcinoma tumour.
XX
OS  Homo sapiens.
XX
PN  WO200157270-A2.
XX
PD  09-AUG-2001.
XX
PE  29-JAN-2001; 2001WO-US00661.
XX
PR  04-FEB-2000; 2000US-0180312.
PR  26-MAY-2000; 2000US-0207456.
PR  30-JUN-2000; 2000US-0608408.
PR  03-AUG-2000; 2000US-0632366.
PR  21-SEP-2000; 2000US-0234687.
PR  27-SEP-2000; 2000US-0236359.
PR  04-OCT-2000; 2000GB-0024263.
XX
PA  (MOLE-) MOLECULAR DYNAMICS INC.
XX
PT  Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR  WPI: 2001-476286/51.
XX
PT  Novel single exon nucleic acid probe used to measuring gene expression
XX  in a human breast -
XX
PS  Claim 25; SEQ ID NO 7731; 322pp; English.
XX
CC  The present invention relates to novel single exon nucleic acid probes.
CC  The present sequence is one such probe. The probes are useful for
CC  measuring human gene expression in a human breast sample, where the probe
CC  hybridises at high stringency to a nucleic acid expressed in the human
CC  breast. The probes are useful for predicting, diagnosing, grading,
CC  staging, monitoring and prognosing diseases of the human breast,
CC  particularly those diseases with polygenic aetiology. The diseases
CC  include: breast cancer, disorders of development, inflammatory diseases
CC  of the breast, fibrocystic changes, proliferative breast disease and
CC  non-carcinoma tumours.
CC  Note: The sequence data for this patent did not form part of the printed
CC  specification, but was obtained in electronic format directly from WIPO
CC  at ftp.wipo.int/pub/published_pct_sequences.
XX
XX  Sequence 830 BP; 193 A; 86 C; 336 G; 215 T; 0 other;

alignment_scores:
      Quality: 180.50      Length: 101
      Ratio: 3.167      Caps: 3
Percent Similarity: 56.436      Percent Identity: 45.545

alignment_block:
US-09-554-547-16 x AA107740
```

```
Align seg 1/1 to: AA107740 from: 1 to: 830
39 TyrAlaGlyIleLeuArgAlaGlyProArgGlyTyrProGlyLeuThrAl 55
      ||:::|||||
236 TATGAGGTGAGAGGTGAGGTGAGTATGAGGTATGAGTGTGTGTGTGT 277
55 aSerIleGlyGlyValGlyValAlaArgLeuGlyGlyArgAlaGlyValG 72
      ||:::|||||
278 ...GGTGGGTGATGAGAGTATGAGTGTGTGTGTGTGTGTGTGTGTGT 323
72 lYValSerSerTyrGlyTyrProSerTyrProGlyTyrProTyrGly 88
      |||
324 GA.....GGTATGAGTGTGAGAGTGTGTGTGTGTGTGTGTGTGTGT 361
89 GlyTyrGlyGlyTyrGlyGlyTyrGlyGlyTyrGlyGlyTyrAspGln 105
      |||
362 GGTGGGTGAGTGTGAGTGTATGAGTGTGTGTGTGTGTGTGTGTGTGT 411
105 yPheGlySerAlaTyrGlyGlyTyrProGlyTyrGlyTyrTyrTyrP 122
      ||:::|||||
412 TTTATGCTGCTGCTGCTGCTGT.....G 434
122 roserglytyrglyglytyrglytyrglyserglyserglysergly 138
      ::::::::::::::::::::
435 ATGAGCTATGCTGCTGCTGCTGCTGCTGCTATGCTGCTGCTGCTGCT 484
139 Gly 139
      |||
485 GGT 487

seq_name: /SID58/gcdata/geneseq/geneseqn/NA2000.DAT:AAC33029

seq_documentation_block:
ID  AAC33029 standard; DNA: 1423 BP.
XX
AC  AAC33029;
XX
DT  17-OCT-2000 (first entry)
XX
DE  Arabidopsis thaliana DNA fragment SEQ ID NO: 1526.
XX
KW  Hybridisation assay; genetic mapping; gene expression control;
XX  protein identification; signal transduction pathway;
XX  metabolic pathway; promoter; termination sequence; ss.
XX
OS  Arabidopsis thaliana.
XX
PN  EP1033405-A2.
XX
PD  06-SEP-2000.
XX
PE  25-FEB-2000; 2000EP-0301439.
XX
PR  25-FEB-1999; 99US-0121825.
PR  05-MAR-1999; 99US-0123180.
PR  09-MAR-1999; 99US-0123548.
PR  23-MAR-1999; 99US-0125788.
PR  25-MAR-1999; 99US-0126264.
PR  29-MAR-1999; 99US-0126785.
PR  01-APR-1999; 99US-0127462.
PR  06-APR-1999; 99US-0128234.
PR  08-APR-1999; 99US-0128714.
PR  16-APR-1999; 99US-0129845.
PR  19-APR-1999; 99US-0130077.
PR  21-APR-1999; 99US-0130449.
PR  23-APR-1999; 99US-0130510.
PR  28-APR-1999; 99US-0130891.
PR  30-APR-1999; 99US-0131449.
PR  30-APR-1999; 99US-0132048.
PR  04-MAY-1999; 99US-0132407.
PR  05-MAY-1999; 99US-0132484.
PR  05-MAY-1999; 99US-0132485.
```

PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.

PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147393.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 23-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.

PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160880.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 23-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

alignment_scores:
 Quality: 180.50 Length: 141
 Ratio: 2.375 Gaps: 11
 Percent Similarity: 53.901 Percent Identity: 41.844

alignment_block:
 US-09-554-547-16 x AAC33029 ..

Align seg 1/1 to: AAC33029 from: 1 to: 1423

33 ArgValHisGly.....AsnLeuTyrAlaGlyIleGluArgAlaG1 46
 ||| |||||
 793 AGGTCAATGTGAGAGTCTCTAGTATGCTGCG..... 828
 46 yProArgIleTyrProGlyLeuThrAlaSerIleGlyGluValG1A 63
 |||||
 829GCATATGGGTATGCTGCTGACGTAGCGCCCTGAATTGGA. 870
 63 IaaArgLeuGlyArgAlaGlyValGlyAlaSerIleGlyGlyTyrGly 79
 |||||
 871GGAGGTATTAATTAATTATGGGGGTGTAAGTAAGA...GGT 909
 80 Tyr.....ProSerTyrGlyTyr.....ProTyrG1 88
 ||| ||| |||||
 910 TACCGAATGACACACCCCTGGCTATTCATGCTTTTGGACCCATATG 959
 88 y.....GlyTyrGly.....GlyTyrGlyGlyTyrGlyTyrG 100
 ||| |||||
 960 GAGTGGGTTGGTGGTGGAGGATATGCGCGTGGAGAGAGACCATTT 1009
 100 IyGlyTyrAspGlnGlyPheGlySerAlaTyrGlyGlyTyr.....Pro 114
 |||||
 1010 TGGGATATCTCTCGTGGAGAGGGGAGGCTATGGGGATATGTTGGCCT 1059
 115 GlyTyrTyrGlyTyrTyrTyrProSerGlyTyrGlyGlyGlyTyr..... 129
 ||||| ||| ||| |||||
 1060 GGTATACGGCGGTGATATGATCTGCTGGCCAGGTGCGCATATGAGG 1109
 130GlyGlySerTyrGlyGlySerTyrGlyGlySerTyrThTyr...P 144
 ||||| ||||| |||||
 1110 AGCAGGCTGCTCATATGCGAGGCGTATAGTACAGTACATACCATC 1159
 144 roAsnValArgAlaSerAlaGly 151
 |||
 1160 CGTATGCAAGGTAGGTCGAGGT 1182

seq_name: /SID8/gcdata/geneseq/geneseqn/NA2000.DAT:AAF18282

seq_documentation_block:

ID AAF18282 standard; DNA; 1089 BP.

AC AAF18282;

DT 14-MAR-2001 (first entry)

DE Lung cancer associated polynucleotide sequence SEQ ID 301.

XX

KW Human; lung cancer associated protein; neuroprotective; cytostatic;
 KW cardiotoxic; immunomodulatory; muscular active; vlnerary;
 KW gastrointestinal; nephrotoxic; antiinfective; gynecological;
 KW antibacterial; diagnosis; neural disorder; immune disorder; reproductive;
 KW proliferative disorder; wound healing; infectious disease; ds.

XX Homo sapiens.

OS WO20005180-A2.

PN 21-SEP-2000.

PD 08-MAR-2000; 2000WO-US05918.

PF 12-MAR-1999; 99US-0124270.

PR (HUMA-) HUMAN GENOME SCT INC.

PA (ROSE/) ROSEN C A.

PI Ruben SM;

DR MPI: 2000-587514/55.

DR P-PSDB: AAB58406.

PT Lung cancer associated gene sequences, referred to as lung cancer

PS Claim 1; Page 759; 1425pp; English.

XX Polynucleotide sequences AAF17982 - AAF18424 encode human lung cancer
 CC associated proteins represented in AAB58106 - AAB58548. Lung cancer
 CC associated proteins and polynucleotide sequences, their agonists, and
 CC antagonists may have neuroprotective; cytostatic; cardioactive;
 CC immunomodulatory; muscular active general; vulnerary; gastrointestinal
 CC general; nephrotoxic; antiinfective; gynecological; or antibacterial
 CC activity. The invention also includes antibodies specific for the
 CC protein or polynucleotide sequences. The lung cancer associated
 CC polynucleotide sequences may be used for detection of lung cancer,
 CC chromosome identification, as chromosome markers, and for numerous other
 CC diagnostic or research purposes. The proteins may be used to treat
 CC disorders such as neural, immune, muscular, reproductive,
 CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
 CC disorders. The proteins may also be used in the treatment of wounds and
 CC infectious diseases. Polynucleotide sequences AAF18425 - AAF18433 and
 CC peptide AAB58549 are used in the course of the invention for the
 CC identification and characterisation of the polynucleotide and protein
 CC sequences.

XX Sequence 1089 BP; 289 A; 183 C; 294 G; 316 T; 7 other;

alignment_scores:

Quality: 177.50 Length: 148
 Ratio: 2.219 Gaps: 10
 Percent Similarity: 54.054 Percent Identity: 40.541

alignment_block:
 US-09-554-547-16 x AAF18282 ..

Align seg 1/1 to: AAF18282 from: 1 to: 1089

14 AlaLeuThrAsnAlaAlaArgAlaGlyArgLeuGlySerAspLeuAspTh 30
 ||||| ||| |||||
 54 GCATTAGCAATTTCCCAAGGCACTCTTTAAGCGACGTCTCTCTATC 103
 30 rPheGlyArgValHisGlyAsn.....LeuTyrAla...G 41
 ||||| |||||
 104 CCTTGA.....GCCCATCAGTAAATGTTTCCATCTATAGCAGAG 144
 41 IyIleGluArgAlaGlyPProArgIleTyrProGlyLeuThrAlaSerIle 57
 ||| ||||| |||||
 145 GTAGTTATGAGGAGGTGATGATGATTAATGGA.....TTT 182

```

CC to the conditional female sterile plant, and (d) producing hybrid
CC seed. Viable seed formation is prevented as a result of the
CC conversion of protoxin to toxin in the female reproductive
CC structures, and pollen production is prevented on the male sterile
CC plant, thus allowing interplanting of the 2 parents of the hybrid
CC cross in order to provide more efficient pollen transfer. Also
CC provided are expression cassettes, transformed plants, and novel
CC female preferential promoters (see also AAV57371, AAV57377 and
CC AAV57380).
XX
XX Sequence 772 BP; 202 A; 135 C; 263 G; 172 T; 0 other;
SQ
alignment_scores:
    Quality: 175.50      Length: 142
    Ratio: 2.472        Gaps: 7
    Percent Similarity: 50.000    Percent Identity: 38.732
alignment_block:
US-09-554-547-16 x AAV57370 ..
Align seg 1/1 to: AAV57370 from: 1 to: 772
47 ProAarglyTyrrProglyleuThr.....Al 55
   ||| ||| ||||| |||||
160 CCTCCAGGAGAGACCTGGCGCTCAAGATGAGAGTGGTTGGTGGATG 209
   55 aserTlGlglyGlyIvalGlyAlaArgleuGlyGlyArgAlaGlyValG 72
   ::: ||||| ||| ::: ||| ||| ||| |||
210 CAACACGATGTGGAGGGATGGAACACACGACGACGATACGGCGCGAG 259
   72 lyValSerSerTyrrglyTyrrgly.....TyrrProSerTrrp 83
   || |||||: ||||| ||| |||||: |||: |||: |||: |||:
260 GAACACACCAACCTGGATGATGCGCGCGAGGAACACGATGACCGCGATAC 309
   84 Gly.....TyrrProTyrrglyGlyTyrr..... 90
   ||| |||: ||||| ||| ||||| ||| ||||| ||| ||||| |||
310 GGTGGAGGATACAGCGCCATCACCTGGTGGCGGCTACGCGGTGGACA 359
   91 .....GlyGlyTyrrGlyGlyTyrrGlyGlyTyrr 99
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
360 AGGAGGGCCTGGATGTGGATGTGGAGAGGATGTGGAGGTGGCAATGTA 409
   99 yrrGlyGlyTyrr.....AspGinglyPheGlySerAlaTyrrglyGlyTyrr 113
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
410 GTCCGTGGTACGGCGATGACAAATGGTGGTGGCAGTGGCAGTGGGCGGA 459
   114 ProGlyTyrrTyrrglyTyrrTyrrProSerGlyTyrrglyGlyGlyTyrrG 130
   ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| |||
460 AATGCGAATGCTGCT.....GGTACGAGAGAGAGAGAG 494
   130 yGlySerTyrrGlyGlySerTyrrGly...GlySerTyrrThrTyrrProAsn 146
   |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
495 CGGGCGGTATGAGAGGGCGCTACGGCAGTGTGTGTGTGTGTGTGTGTGTGT 544
   146 alArgAlaSerAlaGlyAlaAlaAla 154
   |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
545 GCGGATATCATGGCGGTGGTGGTGGCA 570
seq_name: /STD58/gcgdata/geneseq/geneseq/NA2000.DAT:AAc4494
seq_documentation_block:
ID AAC4494 standard; DNA; 993 BP.
XX
XX AAC4494:
DT 18-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana DNA fragment SPQ ID NO: 44901.
XX
XX Hybridisation assay; genetic mapping; gene expression control;
XX protein identification; signal transduction pathway;
XX metabolic pathway; promoter; termination sequence; ss.
XX

```

XX	Arabidopsis thaliana.	
OS	EP1033405-A2.	
XX		
XX		
PN		
PD	06-SEP-2000.	
XX		
PF	25-FEB-2000; 2000EP-0301439	
XX		
PR	25-FEB-1999; 99US-0121825	
PR	05-MAR-1999; 99US-0123180	
PR	09-MAR-1999; 99US-0125348	
PR	23-MAR-1999; 99US-0125788	
PR	23-MAR-1999; 99US-0126764	
PR	23-MAR-1999; 99US-0126764	
PR	01-APR-1999; 99US-0127462	
PR	06-APR-1999; 99US-0128374	
PR	08-APR-1999; 99US-0128714	
PR	16-APR-1999; 99US-0129845	
PR	15-APR-1999; 99US-0130077	
PR	21-APR-1999; 99US-0130449	
PR	23-APR-1999; 99US-0130510	
PR	28-APR-1999; 99US-0130891	
PR	30-APR-1999; 99US-0132048	
PR	30-APR-1999; 99US-0132407	
PR	04-MAY-1999; 99US-0132484	
PR	05-MAY-1999; 99US-0132485	
PR	05-MAY-1999; 99US-0132486	
PR	06-MAY-1999; 99US-0132487	
PR	07-MAY-1999; 99US-0132683	
PR	11-MAY-1999; 99US-0134256	
PR	14-MAY-1999; 99US-0134218	
PR	14-MAY-1999; 99US-0134219	
PR	14-MAY-1999; 99US-0134321	
PR	14-MAY-1999; 99US-0134720	
PR	18-MAY-1999; 99US-0134768	
PR	19-MAY-1999; 99US-0135124	
PR	20-MAY-1999; 99US-0135144	
PR	21-MAY-1999; 99US-0135553	
PR	24-MAY-1999; 99US-0135629	
PR	25-MAY-1999; 99US-0136021	
PR	27-MAY-1999; 99US-0136792	
PR	28-MAY-1999; 99US-0136792	
PR	01-JUN-1999; 99US-0137222	
PR	03-JUN-1999; 99US-0137528	
PR	04-JUN-1999; 99US-0137502	
PR	07-JUN-1999; 99US-0137724	
PR	08-JUN-1999; 99US-0138094	
PR	10-JUN-1999; 99US-0138540	
PR	10-JUN-1999; 99US-0138847	
PR	14-JUN-1999; 99US-0139191	
PR	16-JUN-1999; 99US-0139452	
PR	16-JUN-1999; 99US-0139453	
PR	17-JUN-1999; 99US-0139492	
PR	18-JUN-1999; 99US-0139494	
PR	18-JUN-1999; 99US-0139461	
PR	18-JUN-1999; 99US-0139461	
PR	18-JUN-1999; 99US-0139462	
PR	18-JUN-1999; 99US-0139463	
PR	18-JUN-1999; 99US-0139750	
PR	18-JUN-1999; 99US-0139750	
PR	21-JUN-1999; 99US-0139617	
PR	22-JUN-1999; 99US-0139689	
PR	23-JUN-1999; 99US-0140353	
PR	23-JUN-1999; 99US-0140354	
PR	24-JUN-1999; 99US-0140695	
PR	28-JUN-1999; 99US-0140823	

PR	29-JUN-1999	9905-01403991
PR	30-JUN-1999	9905-0141287
PR	01-JUL-1999	9905-0141842
PR	01-JUL-1999	9905-0142155
PR	02-JUL-1999	9905-0142054
PR	02-JUL-1999	9905-0142390
PR	06-JUL-1999	9905-0142803
PR	08-JUL-1999	9905-0142920
PR	09-JUL-1999	9905-0142920
PR	12-JUL-1999	9905-0143297
PR	13-JUL-1999	9905-0143542
PR	14-JUL-1999	9905-0143654
PR	15-JUL-1999	9905-0144005
PR	16-JUL-1999	9905-0144085
PR	16-JUL-1999	9905-0144086
PR	19-JUL-1999	9905-0144325
PR	19-JUL-1999	9905-0144331
PR	19-JUL-1999	9905-0144332
PR	19-JUL-1999	9905-0144333
PR	21-JUL-1999	9905-0144814
PR	21-JUL-1999	9905-0145086
PR	21-JUL-1999	9905-0145088
PR	21-JUL-1999	9905-0145085
PR	22-JUL-1999	9905-0145087
PR	22-JUL-1999	9905-0145192
PR	22-JUL-1999	9905-0145193
PR	23-JUL-1999	9905-0145145
PR	23-JUL-1999	9905-0145148
PR	23-JUL-1999	9905-0145218
PR	23-JUL-1999	9905-0145224
PR	26-JUL-1999	9905-0145276
PR	27-JUL-1999	9905-0145913
PR	27-JUL-1999	9905-0145918
PR	27-JUL-1999	9905-0145919
PR	28-JUL-1999	9905-0145951
PR	02-AUG-1999	9905-0146586
PR	02-AUG-1999	9905-0146588
PR	02-AUG-1999	9905-0146589
PR	03-AUG-1999	9905-0147038
PR	04-AUG-1999	9905-0147202
PR	04-AUG-1999	9905-0147304
PR	05-AUG-1999	9905-0147190
PR	05-AUG-1999	9905-0147262
PR	06-AUG-1999	9905-0147303
PR	06-AUG-1999	9905-0147416
PR	09-AUG-1999	9905-0147493
PR	09-AUG-1999	9905-0147535
PR	10-AUG-1999	9905-0148171
PR	11-AUG-1999	9905-0148319
PR	12-AUG-1999	9905-0148341
PR	13-AUG-1999	9905-0148565
PR	13-AUG-1999	9905-0148684
PR	16-AUG-1999	9905-0149368
PR	17-AUG-1999	9905-0149125
PR	18-AUG-1999	9905-0149126
PR	20-AUG-1999	9905-0149722
PR	20-AUG-1999	9905-0149723
PR	22-AUG-1999	9905-0149902
PR	23-AUG-1999	9905-0149930
PR	25-AUG-1999	9905-0150566
PR	26-AUG-1999	9905-0150884
PR	27-AUG-1999	9905-0151065
PR	27-AUG-1999	9905-0151066
PR	27-AUG-1999	9905-0151080
PR	30-AUG-1999	9905-0151303
PR	31-AUG-1999	9905-0151538
PR	01-SEP-1999	9905-0151530
PR	07-SEP-1999	9905-0153263
PR	10-SEP-1999	9905-0153070


```

PR 13-SEP-1999: 99US-0153758.
PR 15-SEP-1999: 99US-0154018.
PR 16-SEP-1999: 99US-0154039.
PR 20-SEP-1999: 99US-0154779.
PR 22-SEP-1999: 99US-0155139.
PR 23-SEP-1999: 99US-0155486.
PR 24-SEP-1999: 99US-0155659.
PR 28-SEP-1999: 99US-0156458.
PR 29-SEP-1999: 99US-0156596.
PR 04-OCT-1999: 99US-0157117.
PR 05-OCT-1999: 99US-0157753.
PR 06-OCT-1999: 99US-0157865.
PR 07-OCT-1999: 99US-0158029.
PR 08-OCT-1999: 99US-0158232.
PR 12-OCT-1999: 99US-0158359.
PR 13-OCT-1999: 99US-0159293.
PR 13-OCT-1999: 99US-0159294.
PR 13-OCT-1999: 99US-0159295.
PR 14-OCT-1999: 99US-0159329.
PR 14-OCT-1999: 99US-0159330.
PR 14-OCT-1999: 99US-0159331.
PR 14-OCT-1999: 99US-0159637.
PR 14-OCT-1999: 99US-0159638.
PR 18-OCT-1999: 99US-0159584.
PR 21-OCT-1999: 99US-0160741.
PR 21-OCT-1999: 99US-0160757.
PR 21-OCT-1999: 99US-0160758.
PR 21-OCT-1999: 99US-0160770.
PR 21-OCT-1999: 99US-0160814.
PR 21-OCT-1999: 99US-0160815.
PR 22-OCT-1999: 99US-0160980.
PR 22-OCT-1999: 99US-0160981.
PR 22-OCT-1999: 99US-0160989.
PR 25-OCT-1999: 99US-0161404.
PR 25-OCT-1999: 99US-0161405.
PR 25-OCT-1999: 99US-0161406.
PR 26-OCT-1999: 99US-0161359.
PR 26-OCT-1999: 99US-0161360.
PR 26-OCT-1999: 99US-0161361.
PR 28-OCT-1999: 99US-0161920.
PR 28-OCT-1999: 99US-0161992.
PR 28-OCT-1999: 99US-0161993.
PR 29-OCT-1999: 99US-0162142.

```

alignment_scores:

```

Quality: 175.50      Length: 147
Ratio: 2.372         Gaps: 6
Percent Similarity: 50.340      Percent Identity: 36.735

```

alignment_block:

```
US-09-554-547-16 x AAC44994 ..
```

```
Align seg 1/1 to: AAC44994 from: 1 to: 993
```

```

2 LysAlaPhePheValLeuSerLeuSerThrAlaAlaLeuThrAsnAl 18
   :::::||||| :::::||||| :::::||||| :::::|||||
304 AGCGATTTGGATTGTTGTTACCTTTCACGATTCAGACGCGATCTAGTGC 353
   18 AAlaArgAla.....GlyArgLeuGlySerAspLeuAspThrPheGlyA 33
   I :::::||||| :::::||||| :::::|||||
354 TATCCAGCGCTTTCGATGCGCGC.....GATTCGACATGCGC 388
   33 rGValHLeGlyAsnLeuPyrGAlaGlyIleGlyAlaGAlaLeuProArgGly 49
   ||||| :::::||||| :::::|||||
389 GGGTGTCTTAACATAACTATGCAAAATGATAGAACAAAGTGT..... 429
   50 TyrProGlyLeuThrAlaSerIleGlyGlyValAlaGlyAlaArgLeuGl 66
   ||||| :::::||||| :::::|||||
430 .....GGTGTCTTTCGAGGTGGTGGCGCTATG 455
   66 yGlyArgAlaGlyValAlaGlyValSerSerTyrGlyTyrGlyTyrProSerT 83
   ||||| :::::||||| :::::|||||
456 AGCGCGTGTGCTGTGTTATGCTGTAGTGTGCGTACGCA..... 495

```

```
seq_name: /SIDS8/gcdata/geneseq/geneseqn/NA2001.DAT:AA158656
```

seq_documentation_block:

```

ID AA158656 standard; cDNA: 1505 BP.
XX
AC AA158656;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 859.
XX
KW Human; nootropic; immunosuppressant; cytosstatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokine; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia; ss.
XX
OS Homo sapiens.
XX
PN M0200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000OWO-US34263.
XX
PR 21-JAN-2000; 2000OUS-0488725.
PR 25-APR-2000; 2000OUS-0552317.
PR 09-JUL-2000; 2000OUS-05586042.
PR 19-JUL-2000; 2000OUS-0620312.
PR 03-AUG-2000; 2000OUS-0653450.
PR 14-SEP-2000; 2000OUS-0662191.
PR 19-OCT-2000; 2000OUS-0693036.
PR 29-NOV-2000; 2000OUS-0727344.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Wang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Tang JT, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
WP: 2001-442253/47.
DR P-PSDB; AAM39500.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX
PS Claim 1; SEQ ID NO 859; 10078bp; English.
XX
CC The invention relates to human nucleic acids (AA157798-AA161369) and
CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cytosstatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as

```

CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilization of the activities such as: immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokine activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.

XX Sequence 1505 BP: 425 A; 225 C; 470 G; 385 T; 0 other:

alignment_scores:

Quality: 175.50 Length: 152
 Ratio: 2.250 Gaps: 10
 Percent Similarity: 51.316 Percent Identity: 37.500

alignment_block:

US-09-554-547-16 x AAI58656 ..

Align seg 1/1 to: AAI58656 from: 1 to: 1505

```

20  AYAALAGLYARGLYGLYSEASPLEUASPTHPHEGLYARGYALHISGL 36
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
747  CGCGGAGGAACTTGGAGGTGGTGAATTTGGCCGT...GCTGG 793
    |||:|||||:|||||:|||||:|||||:|||||:|||||:
36  YASNLEU..... 38
    |||:|||||:|||||:|||||:|||||:|||||:
794  AACTTTGGTGAAGAGGAGCTATGCTGAGGCTGGGACGACGAG 843
    |||:|||||:|||||:|||||:|||||:|||||:|||||:
39  ....TYRAGLYLLEGLUARGALAGLYPROARGLYTYRPROGLYLEU 53
    |||:|||||:|||||:|||||:|||||:|||||:|||||:
844  GTAGTATGAGAGA.....GCTGATGCTGATATATGCA... 878
    |||:|||||:|||||:|||||:|||||:|||||:|||||:
54  THALASERLEGLYGLYVALAGLYALARGLEUGLYLARGALAGL 70
    |||:|||||:|||||:|||||:|||||:|||||:|||||:
879  ....TTTGAGCTGAT...GGTGGCACTATGCGCGT.....GG 910
    |||:|||||:|||||:|||||:|||||:|||||:|||||:
70  YVALGLYALSERSER.....TYRGLYTYRGLYTYRPROSERTRPG 84
    |||:|||||:|||||:|||||:|||||:|||||:|||||:
911  TCCTGCTATAGTAGTAGAGGCGCTATGCTGCTGGACCGAGATATG 960
    |||:|||||:|||||:|||||:|||||:|||||:|||||:
84  LYTyrProTYrGLYGLYTYrGLYGLYTYrGLYTYrGLYTYr... 99
    |||:|||||:|||||:|||||:|||||:|||||:|||||:
961  GAACCAAGGTGGTATGATGAGAGTGGAGATATGATGATTACAT 1010
    |||:|||||:|||||:|||||:|||||:|||||:|||||:
100  .....GLYGLYTYRASPGLNGLYPHEGLYSERATATYRGL 111
    |||:|||||:|||||:|||||:|||||:|||||:|||||:
1011  GAAGAGGAAATTTGGCGTGTACTATGCTGTGGGAACTATATA 1060
    |||:|||||:|||||:|||||:|||||:|||||:|||||:
111  YGLYTYRPROGLYTYRGLYTYRGLYTYRPROSERGLYTYRGLYGLY 128
    |||:|||||:|||||:|||||:|||||:|||||:|||||:
1061  TGATTTTGAATATATAGTGA...CAACAGCAATCAATATATAGACCA 1107
    |||:|||||:|||||:|||||:|||||:|||||:|||||:
128  LYTyrGLYGLYSerTYrGLYGLY.....SerTYrGLYGLY 139
    |||:|||||:|||||:|||||:|||||:|||||:|||||:
1108  TGAAGGCGGCACTTTGCTGAAAGACTCGGCACTCCCTATGCTGT 1157
    |||:|||||:|||||:|||||:|||||:|||||:|||||:
140  SerTYr 141
    |||:|||||:|||||:|||||:|||||:|||||:|||||:
1158  GGTAT 1163
    |||:|||||:|||||:|||||:|||||:|||||:|||||:
seq_name: /SID8/gcdata/geneseq/geneseqn/NA2001.DAT:AAF89925
seq_documentation_block:
ID  AAF89925 standard; DNA; 3077 BP.
XX
XX  AAF89925;
AC
XX
XX  06-AUG-2001 (first entry)
DT
XX
XX  Nucleotide sequence of the corn silk C3 gene.
DE
```

XX Corn silk gene; C3 gene; C3 regulatory region; female sterile plant;
 KW pathogen resistance; silk tissue; pistil tissue; stigma; ss.

OS Zea mays.

XX Key Location/Qualifiers

FT promoter

FT CDS

FT 3'UTR

FT CA2321269-A1.

XX 05-APR-2001.

XX 05-OCT-2000; 2000CA-2321269.

XX 05-OCT-1999; 99US-0412704.

XX (MIMC) CANADA MIN AGRIC & AGRI-FOOD CANADA.

XX Tao T, Singh J, Ouellet T;

XX WPI: 2001-329438/35.

XX P-PSDB: AAB83953.

XX New regulatory regions of corn silk for directing expression of a gene

XX of interest in stigma of both monocot and dicot plants

XX Example 3; Page 45-47; 50pp: English.

XX The present sequence represents a corn silk gene, the C3 gene. The

XX regulatory region of the C3 gene is useful for producing a female sterile

XX plant, or a plant with or improved pathogen resistance. A vector

XX comprising a chimeric gene in association with the regulatory region is

XX useful for producing a plant expressing a gene of interest within silk or

XX pistil tissue. The C3 regulatory region may be used to direct the

XX expression of a gene of interest in the stigma of both monocot and dicot

XX plants. Furthermore, the activity of the regulatory region may be

XX modified by the presence of other regulatory regions, for example,

XX enhancers, core promoter elements etc..

XX Sequence 3077 BP: 897 A; 598 C; 721 G; 861 T; 0 other:

alignment_scores:

Quality: 175.50 Length: 142
 Ratio: 2.472 Gaps: 7
 Percent Similarity: 50.000 Percent Identity: 38.732

alignment_block:

US-09-554-547-16 x AAF89925 ..

Align seg 1/1 to: AAF89925 from: 1 to: 3077

```

47  PROARGLYTYRPROGLYLEU.....THRAL 55
    |||:|||||:|||||:|||||:|||||:|||||:|||||:
2116  CCTGCAGAGAGCGCTGCTCAAGATGAGAGTGTGGTGTAGATA 2165
    |||:|||||:|||||:|||||:|||||:|||||:|||||:
55  ASERTLEGLYGLYVALAGLYALARGLEUGLYLARGALAGLYALG 72
    |||:|||||:|||||:|||||:|||||:|||||:|||||:
2166  CAAGCATGTGGAGGCTATGTAACACACAGCCGCTATGCGCGGAG 2215
    |||:|||||:|||||:|||||:|||||:|||||:|||||:
72  LYValSerSerTYrGLYTYrGLY.....TYRPROSerTRP 83
    |||:|||||:|||||:|||||:|||||:|||||:|||||:
2216  GAACCAAGCAACCTGATACGGCGGCGGAGAAACAGCCCGGATAC 2265
    |||:|||||:|||||:|||||:|||||:|||||:|||||:
84  GLY.....TYRPROTYrGLYGLYTYr..... 90
    |||:|||||:|||||:|||||:|||||:|||||:|||||:
```

```

2266 GGTGGAGGATACAAAGCCCATCACCGTCGCGGCCCTAACGGGCTCGGACA 2315
91 ..... Glycyltyrtylglytyrglytyr 99
2316 AGGAGGCCCTGGATGTGGATCTGGAGGAGGATATGGAGGTGCCAAAAGTGA 2365
99 YRGLYGLTYR.....AspIncllypheglyseralatyrglygltyr 113
2366 GTCTCGGGGTACGGCAGTGCACATGTGTGTGTGCATGTGCCTGGCGGGTGA 2415
114 Proglytyrtyrglytyrtyrtyrtyrtyrtyrtyrtyrtyrtyrtyr 130
2416 AATGGCATTGCTGGT.....GGCTACGGAGAGAGAGAGAG 2450
130 yGlySerTyrgLySerTyrgLy...GlySerTyrtHrTyrProAsny 146
|||||:::|||||:::|||||:::|||||:::|||||:::
2451 CGCGCGTTATGGAGCGCGCTACGGCAGCTGTAGTGCTACAGCAGAGAG 2500
146 eAlArgAlaSeAlaglyAlaIAlaIla 154
2501 GCGGATATCATGGCGCGCTGTGTCCA 2526

seq_name: /SID58/gcgdata/geneseq/geneseqn/NA2001.DAT:AAI60442

seq_documentation_block:
ID   AAI60442 standard; cDNA; 3678 BP.
AC   AAI60442:
XX
XX
XX    22-OCT-2001 (first entry)
XX
DE   Human polynucleotide SEQ ID NO 4431.
XX
XX   Human; neotropic; immunosuppressant; cytostatic; gene therapy; cancer;
XX   peripheral nervous system; neuropathy; central nervous system; CNS;
XX   Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
XX   amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
XX   leukæmia; ss.
XX
OS   Homo sapiens.
PN   WO200153312-A1.
PD   26-JUL-2001.
PF   26-DEC-2000; 2000WO-US34263.
XX
PR   21-JAN-2000; 2000US-0488725.
PR   25-APR-2000; 2000US-0532317.
PR   09-JUL-2000; 2000US-0598042.
PR   19-JUL-2000; 2000US-0620312.
PR   03-AUG-2000; 2000US-0653450.
PR   14-SEP-2000; 2000US-0662191.
PR   19-OCT-2000; 2000US-0693036.
PR   29-NOV-2000; 2000US-0727344.
XX
PA   (HYSE-) HYSEQ INC.
PI   Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D,
PI   Wang J, Wang Z, Wehtman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI   Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
XX   WPI: 2001-442253/47.
XX   P-PsDB; AAM41266.
XX
XX   Novel nucleic acids and polypeptides, useful for treating disorders
XX   such as central nervous system injuries -
XX
XX   Claim 1; SEQ ID NO 4431; 10078bp; English.
XX
XX   The invention relates to human nucleic acids (AAI57798-AAI61369) and
XX   the encoded polypeptides (AAM38642-AAW42213) with neotropic,
XX

```

[illegible]


```
29 prhrphgIyArGValHIsGlyAsnLeuTyraAlaGlyIleGluArgAlaG 46
   ::::: ||| |||::: |||:::
1822 AGGCTACGGGTGTCAGAGGTGA.....TACGGTCCCGAGCAGACAGCTG 1865
   46 lypRoArGgIyTyRProGlyLeuThrAlaSerIleGlyGluValGly 62
   ::::: ||| |||::: |||::: |||:::
1866 GTGCTCGCGGAGCTGCTGA.....GCAGAGCTGAGCGCGCTGGCGGT 1909
   63 AlaArgLeuGlyIyArGValHIsGlyValSerSerTyRgIyTyRgI 79
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1910 TACAGTAGAGGTGTCGTGACAGAGCCGCTGCTGAGACTGAGACCCG 1959
   79 yTyRProSerTyRgIyTyRProTyRgIyTyRgIyTyRgIyTyRgIy 96
   ::::: ||| |||::: |||::: |||::: |||::: |||::: |||:::
1960 TCGAGGTGA...GGAGCAGAGCTGAGAGCTACGGTGTGCAAGGTGAT 2006
   96 yT..... 96
2007 ACCGTCCCGAGCAGAGAGCTGTGCTGCTGACAGCTGCTGTCAGAGATCC 2056
   97 GLyGlyTyRgIyTyRAspGln.....G1 105
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2057 GGAGGCGCTGCTGCTACGTACAGGTGCTGCTGCTGAGACCCCTGCAGG 2106
   105 yPheGlySerAlaTyRgIyTyRProGlyTyRgIyTyRgIyTyRgIyTyR 122
   |||::: |||::: |||::: |||::: |||::: |||::: |||:::
2107 AGCTGAGCGGCTGACAGTGTGACAGCAGAGAGCTGAGGCTACGGTGTG 2156
   122 roSerGlyTyRgIy..... 126
2157 AAGGTGATACGTGCGGAGCAGAGCTGCTGACAGCTGCTGTCAGAGAGA 2206
   127 .....GlyGlyTyRgIy...GlySerTyRgIyGlySerTyRgIyG1 139
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2207 GCCGAGCTGAGGTACGAGAGAGGTGCTGCTGCTGAGAGCTACGGTGG 2256
   139 ySerTyRThTyRProAsnValArgAlaSerAlaGlyAlaAla 154
   ::::: ||| ::::: |||::: |||::: |||::: |||::: |||:::
2257 ACAGAGAGGATAT...GTGCTCGAGCTGAGAGCGGTGCTGCTGCA 2299

seq_name: /cgn2_6/ptodata/2/lna/5A_COMB.seq:US-08-458-298-1

seq_documentation_block:
; Sequence 1, Application US/08458298
; Patent No. 5756677
; GENERAL INFORMATION:
; APPLICANT: Lewis, Randolph V.
; TITLE OF INVENTION: cDNAs Encoding Minor Ampullate Spider
; TITLE OF INVENTION: Silk Proteins
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch
; STREET: P.O. Box 747
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22040-3487
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/458,298
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/209,747
; FILING DATE: 14-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Murphy Jr., Gerald M.
; REGISTRATION NUMBER: 28,977
```

```
REFERENCE/DOCKET NUMBER: 1447-104P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-205-8000
TELEFAX: 703-205-8050
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2793 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Nephila clavipes
TISSUE TYPE: minor ampullate gland
FEATURE:
NAME/KEY: CDS
LOCATION: 183..2675
OTHER INFORMATION: /product="N. clavipes minor
OTHER INFORMATION: ampullate silk protein"
```

US-08-458-298-1

```
alignment_scores:
Quality: 172.00 Length: 182
Ratio: 1.911 Gaps: 8
Percent Similarity: 49.451 Percent Identity: 34.066
```

alignment_block:

US-09-354-547-16 x US-08-458-298-1 ..

Align seg 1/1 to: US-08-458-298-1 from: 1 to: 2793

```
13 AlaAlaLeuThrsnAlaAlaArgAlaGlySerAspLeuAs 29
   ||||| ::::: |||::: |||::: |||::: |||::: |||::: |||:::
1772 GCTGCAAGTCTGAGCTGAGCGCTGCGAGGTGCAAGAGCAGAGAGCTGG 1821
   29 prhrphgIyArGValHIsGlyAsnLeuTyraAlaGlyIleGluArgAlaG 46
   |||::: |||::: |||::: |||::: |||::: |||::: |||:::
1822 AGGCTACGGGTGTCAGAGGTGA.....TACGGTCCCGAGCAGACAGCTG 1865
   46 lypRoArGgIyTyRProGlyLeuThrAlaSerIleGlyGluValGly 62
   ::::: ||| ::::: |||::: |||::: |||::: |||::: |||:::
1866 GTGCTCGCGGAGCTGCTGA.....GCAGAGCTGAGCGCGCTGGCGGT 1909
   63 AlaArgLeuGlyIyArGValHIsGlyValSerSerTyRgIyTyRgIy 79
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1910 TACAGTAGAGGTGTCGTGACAGAGCCGCTGCTGAGACTGAGACCCG 1959
   79 yTyRProSerTyRgIyTyRProTyRgIyTyRgIyTyRgIyTyRgIyTyR 96
   ::::: ||| |||::: |||::: |||::: |||::: |||::: |||:::
1960 TCGAGGTGA...GGAGCAGAGCTGAGAGCTACGGTGTGCAAGGTGAT 2006
   96 yT..... 96
2007 ACCGTCCCGAGCAGAGAGCTGTGCTGCTGACAGCTGCTGTCAGAGATCC 2056
   97 GLyGlyTyRgIyTyRAspGln.....G1 105
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2057 GGAGGCGCTGCTGCTACGTACAGGTGCTGCTGCTGAGACCCCTGCAGG 2106
   105 yPheGlySerAlaTyRgIyTyRProGlyTyRgIyTyRgIyTyRgIyTyRgIy 122
   |||::: |||::: |||::: |||::: |||::: |||::: |||:::
2107 AGCTGAGCGGCTGACAGTGTGACAGCAGAGAGCTGAGGCTACGGTGTG 2156
   122 roSerGlyTyRgIy..... 126
2157 AAGGTGATACGTGCGGAGCAGAGAGCTGCTGACAGCTGCTGTCAGAGAGA 2206
   127 .....GlyGlyTyRgIy...GlySerTyRgIyGlySerTyRgIyG1 139
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2207 GCCGAGCTGAGGTACGAGAGAGGTGCTGCTGCTGAGAGCTACGGTGG 2256
```

```

139 yserTyrThrTyrProGluValArgAlaSerAlaIleAlaAla 154
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
2257 ACAAGACCATAT...GGTCCGAGCTGGAGCCGCTGCTGCTGCA 2299

seq_name: /cgn2_6/plodata/2/1na/6A_COMB.seq:US-08-963-168C-3

seq_documentation_block:
; Sequence 3, Application US/08963168C
; Patent No. 6127166
; GENERAL INFORMATION:
; APPLICANT: Bayley, Hagan
; APPLICANT: Cao, Qunping
; APPLICANT: Wang, Yunjaun
; TITLE OF INVENTION: MOLLUSCAN LIGAMENT POLYPEPTIDES
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/963.168C
; FILING DATE: 03-NOV-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Fasse, Peter J.
; REGISTRATION NUMBER: 32,983
; REFERENCE/DOCKET NUMBER: 07917/059001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 381 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 1...378
; US-08-963-168C-3

Alignment_scores:
      Quality: 171.00      Length: 140
      Ratio: 2.085      Gaps: 7
Percent Similarity: 58.571      Percent Identity: 36.429

Alignment_block:
US-09-554-547-16 x US-08-963-168C-3  ..

Align seg 1/1 to: US-08-963-168C-3 from: 1 to: 381

1 MetLysAlaPhePheValLeuSerLeuLeuSerThrAlaAlaLeuThrAs 17
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
1 ATGAACGCTTACATCTGCTTTCTGCTTCTTATTCGCTGCTGCTGCGCC 50
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
17 nAlaAlaLeuArgAlaGlyArgLeuGlySerAspLeuAspThrPheGlyArgV 34
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
51 CCGCGGATGAGGAGGTGCTGCGGAGT...ATGGCGCTACCGGAGAA 97
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
34 aLHisGlyLysLeuTyrAlaGlyIleGlyAlaArgAlaGlyProArgGlyTyr 50
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
98 TGGAGGCGCGGATGAACGAGGC.....GGATTTC 126

```

```

51 ProGlyLeuThrAlaSerIleGlyGlyValGlyAlaArgLeuGlyG 67
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
127 GCGGATGCGCGGAGGAATGGCGAGGTAAAGCGGA...TTGCGCGG 173
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
67 yArgAlaGlyValGlyValSerSerTyrGlyTyrGlyTyrProSerTrpG 84
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
174 AATGGCGGATTCGGCGCATGGAGTGGAAATGGGC.....G 211
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
84 LyTyrProTyrGlyGlyTyrGlyGlyTyrGlyGlyTyrGlyGly 100
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
212 GAGGTCCA...GGCGATTCGGTGGAAATGGAGGTTTCGAGGAATGGCT 258
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
101 GlyTyrAspGlyGlyPheGlySerAlaTyrGlyGlyTyrProGlyTyrTy 117
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
259 GCGGGAAGAGTGATTCGAGGAATGGCAGTGTGATGGAGGTTTCG 308
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
117 rGlyTyrTyrTyrProSerGlyTyrGlyGlyGlyTyrGlyGlySerTyG 134
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
309 AGGAATG.....GGAGCGGAATGCCGTTTCGGTGGAAATGGCGG 349
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
134 Ly...GlySerTyrGlyGly 139
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
350 GCCAAGTGGATTTGGCGCA 369

seq_name: /cgn2_6/plodata/2/1na/6A_COMB.seq:US-08-963-168C-1

seq_documentation_block:
; Sequence 1, Application US/08963168C
; Patent No. 6127166
; GENERAL INFORMATION:
; APPLICANT: Bayley, Hagan
; APPLICANT: Cao, Qunping
; APPLICANT: Wang, Yunjaun
; TITLE OF INVENTION: MOLLUSCAN LIGAMENT POLYPEPTIDES
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/963.168C
; FILING DATE: 03-NOV-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Fasse, Peter J.
; REGISTRATION NUMBER: 32,983
; REFERENCE/DOCKET NUMBER: 07917/059001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 411 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 1...408
; OTHER INFORMATION: The amino acid translation for SEQ ID NO. 6127166. 1
; US-08-963-168C-1

```

```

; CURRENT APPLICATION DATA:

```

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-07-915-246-1


```
seq_documentation_block:
; Sequence 1, Application US/07915246
; Patent No. 5401836
; GENERAL INFORMATION:
; APPLICANT: Basczynski, Chris L.
; APPLICANT: Falls, Lynne
; APPLICANT: Bellmare, Guy
; APPLICANT: Boivin, Rodolphe
; TITLE OF INVENTION: A BRASSICA REGULATORY SEQUENCE FOR
; TITLE OF INVENTION: ROOT-SPECIFIC OR ROOT-ABUNDANT GENE EXPRESSION
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McAndrews, Held, and Malloy
; STREET: 500 W. Madison St. 34th Floor
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60661
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/915,246
; FILING DATE: 19920716
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Pochopien, Donald J.
; REGISTRATION NUMBER: 32,167
; REFERENCE/DOCKET NUMBER: 91 P 1125
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312 707-8889
; TELEFAX: 312 707-9155
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1505 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEetical: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Brassica napus
; STRAIN: Westar
; DEVELOPMENTAL STAGE: Somatic
; TISSUE TYPE: Root
; US-07-915-246-1

Alignment_scores:
Quality: 169.00 Length: 161
Ratio: 2.036 Gaps: 6
Percent Similarity: 51.553 Percent Identity: 36.646

Alignment_block:
US-09-554-547-16 x US-07-915-246-1 ..
Align seg 1/1 to: US-07-915-246-1 from: 1 to: 1505

2 LysAlaPhePheValLeuSerLeuLeuSerThrAlaAlaLeuThAsnAl 18
||| ::::||| ||| ::::||| ::::||| ::::|||
564 AAANATCTTTCTTCCTATGCTTCTTATAGCTTAAAGGTTATGTTCTGC 613
18 AAlaArgAlaGlyArgLeuGlySerAspLeuAspThrPheGlyArgValH 35
||| ||||| ||| ||||| ::::
614 AAGACGACGA.....CTTCTTCTCCTCATGACCCGACGATGAAGTCG 657
35 tsgly.....AsnLeuTyRAIaGlyLe..... 42
||| ||||| ::::||| |||||
658 CCGGATACGCCGAGAAAGATGATTGACGTGTATGATGATTGAGCTT 707
```

```
43 .....GluArgAlaGlyProArgL 49
708 GATGCTGCTGTGCTGTGAGAGCTGCTGCGGAGAGAGCTGCTGCTG 757
49 YTYrProGlyLeuThrAlaSerIleGlyGlyAluValGlyAlaArgLeuG 66
||| ||| :::: ||||| ||||| |||||
758 TTACGCTGA...GCTGAAGCATTGTGAGAGAGAGAGCGCTGACATG 804
66 lYglAArgAlaGlyAlaGlyValSerSerTYrGlyTYrGlyTYrProser 82
||| ||||| ||| ||| ||| |||||
805 GTGCTGTGCTGCTGAGCT.....GCTGCTGCTGCTGCTGCTGCA 842
83 TrpGlyTYrProTYrGlyGlyTYrGlyGlyTYrGlyTYrGlyTYrGly 99
||| ||||| ||| ||||| ||||| |||||
843 GGAGATCTGCTTATGAGAGCTGGAAGCGGTGAAGGTGTGAGCTGATA 892
99 rGlyGlyTYrAspGln.....GlyPheGlySerAlaTYrGlyG 112
||| ||||| ||| ||||| ||||| |||||
893 CCGAGCGGAGAGAGCTGAGACATGTGAGTGTGAGAGAGCGAGAG 942
112 lYTYrProGlyTYrTYrGlyTYrTYrTYrTYrProSerGlyTYrGlyGly 128
||| ||| ||| ||| :::: ||||| ||||| |||||
943 GTGCTGTGAGAGAGCTGCGCGTCCCATGTGCTGATGATGCTGTGCA 992
129 TYrGlyGlySerTYrGlyGlySerTYrGlyGly 139
||| :::: ||||| ||||| |||||
993 GAAGTGTGCTGCTGAGAGAGATATGAGCT 1025

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-963-168C-5

seq_documentation_block:
; Sequence 5, Application US/08963168C
; Patent No. 6127166
; GENERAL INFORMATION:
; APPLICANT: Bayley, Hagan
; APPLICANT: Cao, Qunjaun
; APPLICANT: Wang, Yunjaun
; TITLE OF INVENTION: MOLDSKAN LIGAMENT POLYPEPTIDES
; TITLE OF INVENTION: AND GENES ENCODING THEM
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/963,168C
; FILING DATE: 03-NOV-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Fasse, Peter J.
; REGISTRATION NUMBER: 32,983
; REFERENCE/DOCKET NUMBER: 07917/059001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 399 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 1...396
```

```
OPERATING SYSTEM: Windows95
SOFTWARE: FastSO for Windows Version 3.00
```

Book

```
: Sequence 1, Application US/0814777
: Patent No. 5914265
: GENERAL INFORMATION:
: APPLICANT: Rood, Dennis R.
: APPLICANT: Rothnagel, Joseph A.
: APPLICANT: Greenhalgh, David A.
: APPLICANT: Yuspa, Stuart H.
: TITLE OF INVENTION: KERATIN K1 EXPRESSION VECTORS
: TITLE OF INVENTION: AND METHODS OF USE
: NUMBER OF SEQUENCES: 5
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: LYON & LYON
: STREET: 611 West Sixth Street
: CITY: Los Angeles
: STATE: California
: COUNTRY: U.S.A.
: ZIP: 90017
: COMPUTER READABLE FORM:
: MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: IBM MS-DOS (Version 5.0)
: SOFTWARE: Wordperfect (Version 5.1)
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/147,777
: FILING DATE:
: CLASSIFICATION: 800
: PRIOR APPLICATION DATA:
: PRIOR APPLICATION DATA: Including application
: PRIOR APPLICATION DATA: described below:
: APPLICATION NUMBER: 07/876,289
: FILING DATE: April 30, 1992
: APPLICATION NUMBER: Unassigned (204/144)
: FILING DATE: October 29, 1993
: ATTORNEY/AGENT INFORMATION:
: NAME: Warburg, Richard J.
: REGISTRATION NUMBER: 32,327
: REFERENCE/DOCKET NUMBER: 204/153
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (213) 489-1600
: TELEFAX: (213) 955-0440
: TELEX: 67-3510
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 10747 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: US-08-147-777-1

Alignment_scores:
  Quality: 156.50      Length: 101
  Ratio: 2.524         Gaps: 3
Percent similarity: 61.386      Percent identity: 43.564

alignment_block:
US-09-554-547-16 x US-08-147-777-1 ..

Align seg 1/1 to: US-08-147-777-1 from: 1 to: 10747

54 ThrAlaSerIleGlyGlyValGlyAlaArgLeuGlyAlaArgAlaG1 70
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
6056 ACCATCATGTGGAGGTGGCGGCGGAGGAGGTGGC...GGCGGTGGCTACGG 6102
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
70 yValGlyValSerSerTyTyTyTyTyTyTyTyTyTyTyTyTyTyTy 87
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
6103 CTCTGGAGCTACGACGTATAGCTCGGAGGTGTAGCTATGCTTCGAGG 6152
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
87 yTgGlyGlyTyTyTyTyTyTyTyTyTyTyTyTyTyTyTyTyTyTyTy 103
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
6153 GTGGCGGCGCGCGCGGTGGCGGTGGCGGTGGCTCC...GGAGGTGGAGC 6199
```

```
104 GInGlyPheGlySerAlaTyTyTyTyTyTyTyTyTyTyTyTyTyTy 120
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
6200 TATGGCTGTGGAGGTGGCGGCGGCGCATGTGACGTACGCTCCGAGG 6249
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
120 rTyTyProSerGlyTyTyTyTyTyTyTyTyTyTyTyTyTyTyTyTy 137
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
6250 CAGCAGTGGGGGCTACAGAGGTGGCTGTGGAGGCGGCGGCGAGCT 6299
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
137 yGlyGlySerTyTyTyTyTyTyTyTyTyTyTyTyTyTyTyTyTyTy 153
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
6300 CTGGCGGC.....CGGCGCTGTGGCGGCGGAGC 6328
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
154 Ala 154
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
6329 TCT 6331

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-452-872-1
seq_documentation_block:
: Sequence 1, Application US/08452872
: Patent No. 6057298
: GENERAL INFORMATION:
: APPLICANT: Rood, Dennis R.
: APPLICANT: Rothnagel, Joseph A.
: APPLICANT: Greenhalgh, David A.
: APPLICANT: Yuspa, Stuart H.
: TITLE OF INVENTION: KERATIN K1 EXPRESSION VECTORS
: TITLE OF INVENTION: AND METHODS OF USE
: NUMBER OF SEQUENCES: 5
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: LYON & LYON
: STREET: 611 West Sixth Street
: CITY: Los Angeles
: STATE: California
: COUNTRY: U.S.A.
: ZIP: 90017
: COMPUTER READABLE FORM:
: MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: IBM MS-DOS (Version 5.0)
: SOFTWARE: Wordperfect (Version 5.1)
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/452,872
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US/08/147,777
: FILING DATE:
: APPLICATION NUMBER: 07/876,289
: FILING DATE: April 30, 1992
: APPLICATION NUMBER: Unassigned (204/144)
: FILING DATE: October 29, 1993
: ATTORNEY/AGENT INFORMATION:
: NAME: Warburg, Richard J.
: REGISTRATION NUMBER: 32,327
: REFERENCE/DOCKET NUMBER: 204/153
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (213) 489-1600
: TELEFAX: (213) 955-0440
: TELEX: 67-3510
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 10747 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: US-08-452-872-1

Alignment_scores:
  Quality: 156.50      Length: 101
  Ratio: 2.524         Gaps: 3
```

Percent Similarity: 61.386 Percent Identity: 43.564

alignment_block:
US-09-554-547-16 x US-08-452-872-1 ..

Align seg 1/1 to: US-08-452-872-1 from: 1 to: 10747

```

54 ThrAlaSerLleGlyGlyGlyValAlaArgLeuGlyGlyArgAlaG1 70
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
6056 ACCATCAGTGGAGGTGGCAGCAGGAGGTGGC...GGCGTGGCTACGG 6102
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
70 yValGlyValSerSerTyrGlyTyrProSerTTPGlyTyrProT 87
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
6103 CTCTGGAGTACGACGATGCTCCGAGGTGCTAGCTATGCTTCGAG 6152
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
87 yrcGlyGlyTyrGlyGlyTyrGlyGlyTyrGlyTyrGlyTyrAsp 103
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
6153 GTGGCGGGCGGGCGCGCTGGCGAGCTATGCTCC...GGAGTGGCAGC 6199
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
104 GlnGlyPheGlySerAlaTyrGlyGlyTyrProGlyTyrTyrGlyTyr 120
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
6200 TATGGCTCTGGAGGTGGCGGGCGGCGCATGGCAGCTACGGCTCCGAG 6249
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
120 rTyrProSerGlyTyrGlyGlyGlyTyrGlyGlySerTyrGlyGly 137
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
6250 CACACAGTGGGGCTACAGAGGTGCTCTGGAGGCGGGCGGCGCAGCT 6299
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
137 yrcGlyGlySerTyrTyrProAsnValArgAlaSerAlaGlyAla 153
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
6300 CTGGCGGC.....CGGGCTCTGGCGGGGAGC 6328
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
154 Ala 154
6329 TCT 6331

```

seq_name: /cgn2_6/ptodata/2/1na/PCTUS_COMB.seq:PCT-US93-03985-1

seq_documentation_block:

Sequence 1, Application PC/TUS9303985

GENERAL INFORMATION:

APPLICANT: Roop, Dennis R.
APPLICANT: Rothnagel, Joseph A.
APPLICANT: Greenhalgh, David A.
APPLICANT: Yuspa, Stuart H.
TITLE OF INVENTION: DEVELOPMENT OF A VECTOR TO TARGET GENE
TITLE OF INVENTION: EXPRESSION TO THE EPIDERMIS OF TRANSGENIC ANIMALS
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fulbright & Jaworski
STREET: 1301 McKinney, Suite 5100
CITY: Houston
STATE: Texas
COUNTRY: U.S.A.
ZIP: 77010-3095

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US93/03985

FILING DATE: 19930428

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Paul, Thomas D. 32,714

REGISTRATION NUMBER: 32,714

REFERENCE/DOCKET NUMBER: D-5478

TELEPHONE: 713/651-5325

TELEFAX: 713/651-5246

TELEX: 762829

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

```

; LENGTH: 10747 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; PCT-US93-03985-1

```

alignment_scores:

Quality: 156.50 Length: 101
Ratio: 2.524 Gaps: 3
Percent Similarity: 61.386 Percent Identity: 43.564

alignment_block:

US-09-554-547-16 x PCT-US93-03985-1 ..

Align seg 1/1 to: PCT-US93-03985-1 from: 1 to: 10747

```

54 ThrAlaSerLleGlyGlyGlyValAlaArgLeuGlyGlyArgAlaG1 70
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
6056 ACCATCAGTGGAGGTGGCAGCAGGAGGTGGC...GGCGTGGCTACGG 6102
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
70 yValGlyValSerSerTyrGlyTyrProSerTTPGlyTyrProT 87
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
6103 CTCTGGAGTACGACGATGCTCCGAGGTGCTAGCTATGCTTCGAG 6152
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
87 yrcGlyGlyTyrGlyGlyTyrGlyGlyTyrGlyTyrGlyTyrAsp 103
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
6153 GTGGCGGGCGGGCGCGCTGGCGAGCTATGCTCC...GGAGTGGCAGC 6199
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
104 GlnGlyPheGlySerAlaTyrGlyGlyTyrProGlyTyrTyrGlyTyr 120
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
6200 TATGGCTCTGGAGGTGGCGGGCGGCGCATGGCAGCTACGGCTCCGAG 6249
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
120 rTyrProSerGlyTyrGlyGlyGlyTyrGlyGlySerTyrGlyGly 137
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
6250 CACACAGTGGGGCTACAGAGGTGCTCTGGAGGCGGGCGGCGCAGCT 6299
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
137 yrcGlyGlySerTyrTyrProAsnValArgAlaSerAlaGlyAla 153
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
6300 CTGGCGGC.....CGGGCTCTGGCGGGGAGC 6328
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
154 Ala 154
6329 TCT 6331

```

seq_name: /cgn2_6/ptodata/2/1na/5B_COMB.seq:US-08-147-777-3

seq_documentation_block:

Sequence 3, Application US/08147777

Patent No. 5914265

GENERAL INFORMATION:

APPLICANT: Roop, Dennis R.
APPLICANT: Rothnagel, Joseph A.
APPLICANT: Greenhalgh, David A.
APPLICANT: Yuspa, Stuart H.
TITLE OF INVENTION: KERATIN K1 EXPRESSION VECTORS
TITLE OF INVENTION: AND METHODS OF USE
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: LYON & LYON
STREET: 611 West Sixth Street
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90017

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage

COMPUTER: IBM PC compatible

OPERATING SYSTEM: IBM MS-DOS (Version 5.0)

SOFTWARE: WordPerfect (Version 5.1)


```
120 rTyrProSerGlyTyrGlyGlyTyrGlyGlySerTyrGlyGlySert 137
      ::::::::::: ||||| ::::::::::: |||||
20482 CAGCACTGGGGGCTACAGAGTGCTCTGAGCGCGCGCGGCGAGCT 20531
137 yrglyglyserTyrThrTyrProAsnValArgAlaSerAlaGlyAlaAla 153
      |||||
20532 CTGGCGGCGC.....CGGGGCTCTGGCGCGCGGAGC 20560
154 Ala 154
20561 TCT 20563

seq_name: /cgn2_6/ptodata/2/1na/PCTUS_COMB.seq:PCT-US93-03985-3
seq_documentation_block:
  Sequence 3, Application PC/TUS9303985
  GENERAL INFORMATION:
  APPLICANT: Roop, Dennis R.
  APPLICANT: Rothnagel, Joseph A.
  APPLICANT: Greenhalgh, David A.
  APPLICANT: Yuspa, Stuart H.
  TITLE OF INVENTION: DEVELOPMENT OF A VECTOR TO TARGET GENE
  NUMBER OF SEQUENCES: 5
  CORRESPONDENCE ADDRESSES:
  ADDRESSEE: Fulbright & Jaworski
  STREET: 1301 McKinney, Suite 5100
  CITY: Houston
  STATE: Texas
  COUNTRY: U.S.A.
  ZIP: 77010-3095
  COMPUTER READABLE FORM:
  MEDIUM TYPE: Floppy disk
  OPERATING SYSTEM: IBM PC compatible
  SOFTWARE: Patent Release #1.0, Version #1.25
  CURRENT APPLICATION DATA:
  APPLICATION NUMBER: PCT/US93/03985
  FILING DATE: 19930428
  CLASSIFICATION:
  ATTORNEY/AGENT INFORMATION:
  NAME: Paul, Thomas D.
  REGISTRATION NUMBER: 32,714
  REFERENCE/DOCKET NUMBER: D-5478
  TELECOMMUNICATION INFORMATION:
  TELEPHONE: 713/651-5325
  TELEFAX: 713/651-5246
  TELEX: 762829
  INFORMATION FOR SEQ ID NO: 3:
  SEQUENCE CHARACTERISTICS:
  LENGTH: 24979 base pairs
  TYPE: NUCLEIC ACID
  STRANDEDNESS: double
  TOPOLOGY: linear
  MOLECULE TYPE: DNA (genomic)
  HYPOTHETICAL: NO
  ANTI-SENSE: NO
  PCT-US93-03985-3

alignment_scores:
  Quality: 156.50      length: 101
  Ratio: 2.524         Gaps: 3
  Percent Similarity: 61.386      Percent Identity: 43.564

alignment_block:
US-09-554-547-16 x PCT-US93-03985-3
Align seg 1/1 to: PCT-US93-03985-3 from: 1 to: 24979
54 ThrAlaSerTleGlyGlyGlyAlaValGlyAlaArgLeuGlyGlyArgAlaG1 70
||| ||| |||||::: ||::: ||||| |||||
```

```
20288 ACCATCAGTGGAGGTGGCAGCCGAGGAGGTGCT...GGCGGTGCTACGG 20334
70 yValGlyAlaSerSerTyrGlyTyrGlyTyrProSerTyrGlyTyrPro 87
      ||| ||||| ||||| ||| |||||
20335 CTCTGAGAGTACAGACTATGCTCCGAGGTGTACTATGTTTCGGAG 20384
87 yrglyglyserTyrGlyGlyTyrGlyGlyTyrGlyGlyTyrGlyTyrAsp 103
      ||||| ||||| ||||| ||||| ||||| |||||
20385 GTGGCGGCGCGCGCGCGCGCTGAGCTATGGCTTC...GGAGGTGGCAGC 20431
104 GInGlyPheGlySerAlaTyrGlyGlyTyrProGlyTyrGlyTyrTyr 120
      ||| ||||| ::::::::::: ||||| ::::::::::: |||||
20432 TATGCTCTGAGAGGTGGCGCGCGCGCCAGCTACGCTCCGAG 20481
120 rTyrProSerGlyTyrGlyGlyTyrGlyGlySerTyrGlyGlySert 137
      ::::::::::: ||||| ::::::::::: |||||
20482 CAGCACTGGGGGCTACAGAGTGCTCTGAGCGCGCGCGGCGAGCT 20531
137 yrglyglyserTyrThrTyrProAsnValArgAlaSerAlaGlyAlaAla 153
      |||||
20532 CTGGCGGCGC.....CGGGGCTCTGGCGCGCGGAGC 20560
154 Ala 154
20561 TCT 20563

seq_name: /cgn2_6/ptodata/2/1na/5B_COMB.seq:US-09-010-9288-1
seq_documentation_block:
  Sequence 1, Application US/090109288
  Patent No. 5994099
  GENERAL INFORMATION:
  APPLICANT: Lewis, Randolph V
  APPLICANT: Hayashi, Cheryl Y
  TITLE OF INVENTION: EXTREMELY ELASTIC SPIDER SILK PROTEIN AND DNA
  NUMBER OF SEQUENCES: 29
  CORRESPONDENCE ADDRESSES:
  ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
  STREET: 8110 GATEHOUSE RD, SUITE 500E
  CITY: FALLS CHURCH
  STATE: VIRGINIA
  COUNTRY: UNITED STATES OF AMERICA
  ZIP: 22042
  COMPUTER READABLE FORM:
  MEDIUM TYPE: Floppy disk
  OPERATING SYSTEM: IBM PC compatible
  SOFTWARE: Patent Release #1.0, Version #1.30
  CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/09/010,928B
  FILING DATE: 22-JAN-1998
  CLASSIFICATION: 435
  ATTORNEY/AGENT INFORMATION:
  NAME: Murphy Jr., Gerald M
  REGISTRATION NUMBER: 289377
  REFERENCE/DOCKET NUMBER: 1447-109P
  INFORMATION FOR SEQ ID NO: 1:
  SEQUENCE CHARACTERISTICS:
  LENGTH: 2830 base pairs
  TYPE: nucleic acid
  STRANDEDNESS: not relevant
  TOPOLOGY: linear
  MOLECULE TYPE: cDNA
  HYPOTHETICAL: NO
  FEATURE:
  NAME/KEY:
  LOCATION: 1..2830
  OTHER INFORMATION: /note= "Flagelliform DNA sequence
  OTHER INFORMATION: taken from the 5' region. The putative start codon is at
  OTHER INFORMATION: position 219"
  FEATURE:
  NAME/KEY: CDS
```

LOCATION: 219..2830
US-09-010-928B-1

alignment_scores:
Quality: 154.50 Length: 138
Ratio: 2.207 Gaps: 7
Percent Similarity: 50.725 Percent Identity: 37.681

alignment_block:
US-09-554-547-16 x US-09-010-928B-1 ..

Align seg 1/1 to: US-09-010-928B-1 from: 1 to: 2830

```
46 GLYPProAGGLYtyr...ProGlyLeuThrAlaSerIleGly...GLYGI 60
   |||||  |||  |||||  |||||  |||||  |||||  |||||  |||||
1935 GGACCTGGTGGCTCCGGACCAAGAGGCGGTAGAGTCTGGTGGCTCCGGACA 1984
   60 uValGlyAlaArgLeuGlyGlyAArgAlaGlyValGlyValSerSerTyrc 77
   : |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
1985 AGGAGGAGTAAGACCTAGTGGCTCCGGACCAAGTGGCGTAGGACTGGAG 2034
   77 LYTYrGlyTYrProSerTrpGlyTYrProTyrcGly..... 88
   || |||  |||  |||  |||  |||  |||  |||  |||
2035 GCGTAGGACCCGCTGGTGGCTGAGACCTTACGGTCTGGTGGTCCGGA 2084
   89 ..... 94
2085 CCCGGAAGTGCAGGAGCGCTGGAGACTTATGACCTGGTGGTTCGG 2134
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
   94 YGlyTYrGlyGlyTYrGlyGlyTYrAAspGlnGlyPheGlySerAlaTyrc 111
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
2135 AGGACCGCGGTGGTTCGAGGA...CCGCGTGGTGGTGGTGGACCTAGC 2181
   111 LY..... 119
   || |||  |||  |||  |||  |||  |||  |||  |||
2182 GTCCAGGTGGTCTGGTGGACCTAGGACCAAGTGGTGGTGGACCC 2231
   120 TYrTYrProSerGlyTYrGlyGlyTYr..... 131
   || |||  |||  |||  |||  |||  |||  |||  |||
2232 TACGGACCAAGTGGTGGTGGACCTAGGCGCGGCGGTGGTGGTGG 2281
   131 YSerTYrGly...GlySerTYrGlyGlySerTYrThrTYrProAsnValA 147
   | |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
2282 ACCCTACGGCGCGGAGGTCGTGGTGGATCTTACGGGCTGGGTGGTGG 2331
   147 rGAlaSerAlaGly 151
   :|||  |||  |||  |||  |||  |||  |||  |||
2332 GTGGATCAGAGAGCT 2345
```

THIS PAGE BLANK (USPTO)

Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:672975

Seq primer: -28m13 rev1 ET from Amersham
 High quality sequence stop: 446.

FEATURES

source

1.494
 /organism="Mus musculus"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone_image="1281175"
 /clone_lib="Stratagene mouse skin (#937313)"
 /sex="females"
 /tissue_type="whole skin"
 /dev_stage="11 weeks old"
 /lab_host="SOLR (kanamycin resistant)"
 /note="Organ: skin; Vector: pBluescript SK-; Site: 1; EcoRI
 : Site: 2; XhoI: Cloned unidirectionally. Primer: Oligo
 dt. Whole skin from 11 week old C57BL/6 female mice.
 Average insert size: 1.0 kb; Uni-ZAP XR Vector; -5'
 adaptor sequence: 5' GAATTCGGCAGG 3' -3' adaptor
 sequence: 5' CTCGAGTTT TTT TTT TTT TTT 3'."
 102 c 165 g 140 t

BASE COUNT

87 a 102 c 165 g 140 t

ORIGIN

alignment_scores:

Quality: 208.50 Length: 112
 Ratio: 2.780 Gaps: 10
 Percent Similarity: 66.964 Percent Identity: 47.321

alignment_block:

US-09-554-547-16 x AA799268 ..

Align seg 1/1 to: AA799268 from: 1 to: 494

```

49 G1YTYrProG1yLeuThra1aSer11eG1yG1uValG1yAla..... 63
||||| ||| :|||:||||| :|||:
34 GCGTAC...GGCTATGGCAGCGCTTTGGAGGC...CTGGCGTGGGATG 77
64 .....ATgLeuG1yG1yArGa1aG1yAlaG1yAlaSer1 76
||||| ||| :||| |||
78 TAACAGCATCCGACACTGGCTGGCTGGCTATGAGCGCTTCGCAT 127
||||| ||| :||| |||
76 YrG1yTYrG1yTYrProSerTrp...G1yTYrProTYrG1y...G1yTYr 90
||||| ||| :||| ||| |||
128 ATGGCTGTGGCTATGAGCGCTTTGGAGCGCTTGGATATGCGCTGCGCAT 177
||||| ||| :||| |||
91 G1yG1yTYrG1y.....G1yTYrG1yG1yTYrG1yTYrAspG1 104
||||| ||| :||| ||| |||
178 GGAGCGTACGAGATACGCTGTGGCTATGAGCGCTTTGGAGCGCTTT 222
||||| ||| :||| |||
104 nG1yPhG1ySerAlaTYrG1y.....G1yTYrProG 115
||||| ||| :||| |||
223 .GGATATGGCTCTGGCTATGAGCGCTTCGCATATGAGCGCTTCGCTATGAG 271
||||| ||| :||| |||
115 1yTYrTYrG1yTYrTYrTYrProSerG1yTYrG1yG1y...G1yTYr 130
||||| ||| :||| ||| ||| |||
272 GCTTTGGAGCGCTTTGGATATGCGCTGTGGCTATGAGCGCTATGAGATATGCC 321
||||| ||| :||| |||
131 G1ySerTYrG1yG1y...SerTYrG1yG1ySerTYr 141
||||| ||| :||| |||
322 TCTGGCTTTGGAGCGCTATGATATGCGCTCTGGCTTC 357
||||| ||| :||| |||

```

seq_name: gb_est1:AA792705

seq_documentation_block:

LOCUS AA792705 578 bp mRNA EST 09-FEB-1998
 DEFINITION vs75d11.r1 Striatagene mouse skin (#937313) Mus musculus cDNA clone
 IMAGE:1152117 5', mRNA sequence.

ACCESSION AA792705
 VERSION AA792705.1 GI:2855660
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus.

REFERENCE

1 (bases 1 to 578)
 Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
 Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
 Teising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
 Waterston, R.

TITLE

The WashU-HMI Mouse EST Project
 Unpublished (1996)

COMMENT

Contact: Marra M/Mouse EST Project
 WashU-HMI Mouse EST Project
 Washington University School of Medicine
 444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:625325

Putative full length read
 vector to vector length is 619
 Seq primer: -28m13 rev1 ET from Amersham
 High quality sequence stop: 439.

FEATURES

source

1.578
 /organism="Mus musculus"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone_image="1152117"
 /clone_lib="Stratagene mouse skin (#937313)"
 /sex="females"
 /tissue_type="whole skin"
 /dev_stage="11 weeks old"
 /lab_host="SOLR (kanamycin resistant)"
 /note="Organ: skin; Vector: pBluescript SK-; Site: 1; EcoRI
 : Site: 2; XhoI: Cloned unidirectionally. Primer: Oligo
 dt. Whole skin from 11 week old C57BL/6 female mice.
 Average insert size: 1.0 kb; Uni-ZAP XR Vector; -5'
 adaptor sequence: 5' GAATTCGGCAGG 3' -3' adaptor
 sequence: 5' CTCGAGTTT TTT TTT TTT TTT 3'."
 115 a 119 c 178 g 166 t

BASE COUNT

115 a 119 c 178 g 166 t

ORIGIN

alignment_scores:
 Quality: 208.50 Length: 112
 Ratio: 2.780 Gaps: 10
 Percent Similarity: 66.964 Percent Identity: 47.321

alignment_block:

US-09-554-547-16 x AA792705 ..

Align seg 1/1 to: AA792705 from: 1 to: 578

```

49 G1YTYrProG1yLeuThra1aSer11eG1yG1uValG1yAla..... 63
||||| ||| :|||:||||| :|||:
22 GCGTAC...GGCTATGGCAGCGCTTTGGAGGC...CTGGCGTGGGATG 65
64 .....ATgLeuG1yG1yArGa1aG1yAlaG1yAlaSer1 76
||||| ||| :||| |||
66 TAACAGCATCCGACACTGGCTGGCTGGCTATGAGCGCTTCGCAT 115
||||| ||| :||| |||
76 YrG1yTYrG1yTYrProSerTrp...G1yTYrProTYrG1y...G1yTYr 90
||||| ||| :||| ||| |||
116 ATGGCTGTGGCTATGAGCGCTTTGGAGCGCTTGGATATGCGCTCTGGCTAT 165
||||| ||| :||| |||
91 G1yG1yTYrG1y.....G1yTYrG1yG1yTYrG1yTYrAspG1 104
||||| ||| :||| |||

```

```

|||||
166 GGAGGCTAGCATGCGCTCTGCTATGAGAGCTTGGAGGCTTT..... 210
104 nglpneglyseralatyrgly.....GLYTYRPro 115
211 .GGATATGCGCTCTGCGCTATGAGAGCTTCGATATGCGCTCGCTATGAGAG 259
115 lYTYRtyrglyTYRtyrProSerGlyTYRtyrgly...GLYTYRgly 130
260 GCTTGGAGGCTTGGATATGCGCTATGAGAGCTATGATATGAGATATGAGC 309
131 GLYserTYrgly...SerTYrglySerTYr 141
310 TCTGGCTTGGAGGCTATGATATGCGCTCGCTTC 345

```

seq_name: gb_est1:AA500430

seq_documentation_block:

LOCUS AA500430 449 bp mRNA EST 01-JUL-1997
DEFINITION v181d03.r1 Stratiagene mouse skin (#937313) Mus musculus cDNA clone
IMAGE:918629 5', mRNA sequence.

ACCESSION AA500430

VERSION KEYWORDS

EST.

SOURCE house mouse.

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 449)
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Gelsel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.

TITLE The WashU-HHMI Mouse EST Project
JOURNAL Unpublished (1996)

COMMENT Contact: Marra M/Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LNL: contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:530845
Seq primer: -28m13 rev1 ET from Amersham
High quality sequence stop: 408.

Location/Qualifiers
1. 449
/organism="Mus musculus"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:918629"
/clone_lib="Stratiagene mouse skin (#937313)"
/sex="females"
/tissue_type="whole skin"
/dev_stage="11 weeks old"
/lab_host="SOLR (kanamycin resistant)"
/note="Organ: skin; Vector: pBluescript SK-; Site:1: EORI
; Site-2: XhoI; Cloned unidirectionally. Primer: Oligo
dT. Whole skin from 11 week old C57BL/6 female mice.
Average insert size: 1.0 kb; Uni-ZAP XR Vector; -5'
adaptor sequence: 5' GAATTCGGCAGAG 3' -3' adaptor
sequence: 5' CTCGAGTTTCTTTTCTTTTCTTTT 3' "

FEATURES
source

BASE COUNT 78 a 88 c 150 g 133 t
ORIGIN

alignment_scores:

Quality: 203.00 Length: 106
Ratio: 2.819 Gaps: 9
Percent Similarity: 67.925 Percent Identity: 47.170

alignment_block:
us-09-554-547-16 x AA500430 ..
Align seg 1/1 to: AA500430 from: 1 to: 449

```

55 AlaserlleglygluvalglyAla.....Argle 65
8 AGCAGCTTGGAGGC...CTGGCGTGGATATGATACAGACATCCCGAGACT 54
65 ugllylAArgAlaglyAlaSerSerTYrglyTYRtyrPro 82
55 GGGCTTGCGCTTGGATATGCGCTTCGGATATGAGGCTTGGAG 104
82 erTP...GLYTYRProTYrgly...GLYTYRglyTYRtyr... 94
105 GCTTGGAGGCTTGGATATGCGCTTCGGATATGAGGCTTGGAG 154
95 ...GLYTYRglyTYRtyrGLYTYRtyrAspGlnGlypneglyseralaty 110
155 TCTGGCTATGAGAGGCTTGGAGGCTT.....GGATATGCGCTCGCTA 198
110 rgly.....GLYTYRProGLYTYRtyrTYRtyr 121
199 TGGAGGCTTCGGATATGCGCTCTGCTATGAGGCTTGGAGCTTGGAT 248
121 yrProSerGLYTYRgly...GLYTYRglySerTYrgly... 135
249 ATGCTTGGCTATGAGGCTATGATATGCGCTTGGAGGCTAT 298
136 SerTYrglySerTYr 141
299 GGATATGCGCTCGCTTC 316

```

seq_name: gb_est1:AA727403

seq_documentation_block:

LOCUS AA727403 469 bp mRNA EST 02-JAN-1998
DEFINITION vU99f12.r1 Stratiagene mouse skin (#937313) Mus musculus cDNA clone
IMAGE:1210319 5', mRNA sequence.

ACCESSION AA727403

VERSION AA727403.1 GI:2745110

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 469)
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Gelsel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.

TITLE The WashU-HHMI Mouse EST Project
JOURNAL Unpublished (1996)

COMMENT Contact: Marra M/Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LNL: contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:646663
Seq primer: -28m13 rev1 ET from Amersham
High quality sequence stop: 441.

FEATURES
source

1. 469
/organism="Mus musculus"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:1210319"

```

/clone_11b="Stratagene mouse skin (#937313)"
/sex="females"
/tissue.type="whole skin"
/dev_stage="11 weeks old"
/lab_host="SOLR (kanamycin resistant)"
/note="Organ: skin; Vector: pBluescript SK-; Site_1: EcoRI
; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo
dr. Whole skin from 11 week old C57BL/6 female mice.
Average insert size: 1.0 kb; Uni-ZAP XR Vector: -5'
adaptor sequence: 5' GATTCGGCAGCAG 3' -3' adaptor
sequence: 5' CTCGAGTTTCTTTTCTTTTCTTTT 3'"

BASE COUNT      78 a      94 c      158 g      139 t
ORIGIN

alignment_scores:
  Quality: 203.00      Length: 113
  Ratio: 2.942      Gaps: 9
  Percent Similarity: 61.062      Percent Identity: 46.018

alignment_block:
US-09-554-547-16 x AA727403 ..

Align seg 1/1 to: AA727403 from: 1 to: 469

53 LeuThrAlaSerIleGlyGluValGlyAlaArgLeuGlyArgAl 69
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
13 CTACCCGAGCGCTGGCTACCGCTATGGCAGCCTTTGGAGGC...CT 59

69 aGlyValGlyValSerSerTyr.....GlyTyrGlyTyrP 81
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
60 GCGCTGTGATGTACAGCATCCCATCTGGCTGTGGCTGCTATGG 109

81 roSerTrpGlyTyrProTyrGlyGlyTyrGlyGlyTyrGly 97
   : : : : : : : : : : : : : : : : : : : : : :
110 TAGGCTTGGCATATGCG...TCTGGCTATGGAGGCTTGGAGC 156

98 .....GlyTyrGlyGlyTyrAspGlnGlyPheGlySerAlaTyrG 111
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
157 TATGGCTCTGGCTATGGAGCGCTAC.....CGATACGCGCTG 200

111 yGlyTyr.....ProGlyTyrTyrGlyTyrTyrTyrP 122
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
201 AGGCTTTGGAGGCTTGGATATGCTGTGGATATGGAGCTTCGGA 250

122 roSerGlyTyrGlyGly.....GlyTyrGlyGlySerTyrGly 134
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
251 GCTGTGGCTATGGAGGCTTGGAGGCTTGGATATGCTGTGGAT 300

135 Gly...SerTyrGly.....GlySerTyrThrTyr 143
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
301 GGCTATGATATGGCTCTGGCTTGGAGGCTATGATAT 339

seq_name: gb_est1:BE196524

seq_documentation_block:
LOCUS      BE196524      962 bp      mRNA      EST      02-MAR-2001
DEFINITION HVSMEH0092116f Hordeum vulgare 5-45 DAP spike EST library
            HVCNDA0009 (5 to 45 DAP) Hordeum vulgare cDNA clone HVSMEH0092116f,
            mRNA sequence.
ACCESSION  BE196524
VERSION    BE196524.1 GI:8708720
KEYWORDS   EST.
SOURCE     barley.
            Hordeum vulgare
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae
            ; Triticeae; Hordeum.
REFERENCE  1 (bases 1 to 962)
            Wing,R., Close,T.J., Kleinholz,A., Wise,R., Begum,D., Fritsch,D., Yu
            Y., Anderson,H., Dale,J., Henry,D., Kernodle,S., Palmer,M., Rambo
            T., Snakl,C., Schwartzbeck,J., Simmons,J., Chol,D.W., Main,D. and
            Wood,T.

```

```

TITLE      Development of a genetically and physically anchored EST resource
            for barley genomics
JOURNAL    Unpublished (2000)
COMMENT    Contact: Wing RA
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Tel: 864 656 7288
            Fax: 864 656 4293
            Email: rwing@clemson.edu
            Seq primer: AATTACCTCCTAAGG
            High quality sequence stop: 723.

FEATURES
  source
    1..962
    /organism="Hordeum vulgare"
    /cultivar="Morex"
    /db_xref="taxon:4513"
    /clone="HVSMEH0092116f"
    /clone_11b="Hordeum vulgare 5-45 DAP spike EST library
    HVCNDA0009 (5 to 45 DAP)"
    /tissue.type="5-45 DAP Spike"
    /lab_host="SOLR"
    /note="Vector: lambdaZAP; Site_1: EcoRI; Site_2: XhoI; For
    more details on library preparation and sequence analysis
    see http://www.genome.clemson.edu/projects/barley/ To
    order a clone see http://www.genome.clemson.edu/orders"

BASE COUNT      242 a      185 c      309 g      224 t      2 others
ORIGIN

alignment_scores:
  Quality: 202.50      Length: 162
  Ratio: 2.355      Gaps: 11
  Percent Similarity: 53.086      Percent Identity: 40.741

alignment_block:
US-09-554-547-16 x BE196524 ..

Align seg 1/1 to: BE196524 from: 1 to: 962

23 ArgLeuGlySerAspLeuAspThrPheGlyArgValHis.....GlyAs 37
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
86 AGGCTATATGTCAGAT.....GGGAGATCATGATCTTGGAGG 123

37 nLeuTyrAlaGlyIleGluArgAlaGlyProArg..... 48
   : : : : : : : : : : : : : : : : : : : : : :
124 GAAGCAGTTGAATMAAAGGCCAGCCAGGAAGAAACCTGGTGCTG 173

49 .....GlyTyrProGlyLeuThrAla 55
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
174 ATTCAAGCTCTACAGTACATAGCATGGAGAGGTGGCCACCGTAG 223

56 SerIleGlyGlyGluValGlyAlaArgLeuGlyGlyArgAlaGly 72
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
224 TCATACCGTGTGGTGAGCGCGCGCAGTGTGTGGAGCGCTTCAT 273

72 yAlaSerSer...TyrGlyTyrGly..... 79
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
274 CAGCAGCAGGCTATGATGATGATGATGATGATGATGATGATGAT 323

80 .....TyrProSerTrpGlyTyrProTyrGlyGlyTyrGly...GlyTyr 93
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
324 CCTACTATGATGATGACAGA.....TATGTGCTCCTACGCTACAG 367

94 GlyGlyTyrGlyGlyTyrGlyGlyTyrAspGlnGlyPheGlySerAla 110
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
368 GGTGCAATATGAGAGTAACCTCCCTATGCGTCAAGTTTGGCTGCT 417

110 rGlyGly.....TyrProGlyTyrTyrGlyTyrTyrTyrProSerGly 125
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
418 TGGTGTCTCATATATATGAGAGCGCGATATGATGATAT..... 454

125 yGlyGlyGlyGlyTyrGlyGlySerTyrGlyGlySerTyrGlyGly 141

```


and propagated in SOLR cells. Clone ordering and sequencing was performed by the Centre National de Sequence (Genoscope, Evry, France)."

BASE COUNT 79 a 93 c 179 g 84 t

ORIGIN

alignment_scores: Quality: 201.00 Length: 112
Ratio: 2.871 Gaps: 7
Percent Similarity: 62.500 Percent Identity: 48.214

alignment_block:
US-09-554-547-16 x AL372330 ..

Align seg 1/1 to: AL372330 from: 1 to: 435

```

31 Phcglarvvalhis.glyasnleuTyraAlaGlylleGluarGalagIyp 47
|||||.....:|||||.....:|||||.....:|||||.....:
19 TTTCCGGCGCTACGCGCGCTACCTCAGTCCGCGCTACGCGCTACGCGCA 68
47 roArGgLyTyrrProGlyLeuThraLaseRIegLyGlyGluValGlyAla 63
69 GCTCGCTATACCT.....GCTTCCTCTACGCGCGCGCTACGCGCGC 112
64 ArgLeuAlGlyArGAlaGlyValGlyValSerSerTyrrGlyTyrrGly 80
|||||.....:|||||.....:|||||.....:|||||.....:
113 .....GTTTCGAGCGCGCGCTACGAGATCCAGCGCGCGCTACGAGG 156
80 rProSerTrpGlyTyrrProTyrrGlyTyrrGly.....TyrrGlyG 95
111 .....:|||||.....:|||||.....:|||||.....:
157 ATCCAGCGCGCGATATGCGGCGGCTATGTCGCGGAGCATCAGCGC 206
95 LyTyrrGlyGlyTyrrGlyTyrrAspGlnGlyPheGlySerAlaTyrrGly 111
|||||.....:|||||.....:|||||.....:|||||.....:
207 GAATATGCGCGA...GGCGTTACGCTGAGGAGCGGCGGAGATATGCT 253
112 GATTyrrProGlyTyrrGlyTyrrTyrrProSerGlyTyrrGlyGly 128
111 .....:|||||.....:|||||.....:|||||.....:
254 GGA.....GGCGGTATATGCT.....GGCGGTATG 276
128 yTyrrGlyGlySerTyrrGlyGlySerTyrrGlyGly 139
|||||.....:|||||.....:|||||.....:|||||.....:
277 CTACGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 310

```

seq_name: gb_est1:AA670823

seq_documentation_block:

LOCUS AA670823 595 bp mRNA EST 25-NOV-1997
DEFINITION vs70G08.r1 StrataGene mouse skin (#937313) Mus musculus cDNA clone
IMAGE:1151630 5', mRNA sequence.

ACCESSION AA670823
VERSION AA670823.1 GI:2642902

KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 595)
AUTHORS Marfa,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Watson,R.

TITLE The WashU-HMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Maria M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu

This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:624838
Seq primer: -28ml3 rev1 ET from Amersham
High quality sequence stop: 447.

FEATURES

source

1..595

/organism="Mus musculus"

/strain="C57BL/6"

/db_xref="taxon:10090"

/clone="IMAGE:1151630"

/clone_id="Stratagene mouse skin (#937313)"

/sex="females"

/tissue_type="whole skin"

/dev_stage="11 weeks old"

/lab_host="SOLR (kanamycin resistant)"

/note="Organ: skin; Vector: pluescript SK-; Site:1; EcorI ; Site:2; XhoI; Cloned unidirectionally. Primer: Oligo dt. Whole skin from 11 week old C57BL/6 female mice. Average insert size: 1.0 kb; Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGACGAG 3' -3' adaptor sequence: 5' CTCGAGTTT TTTT TTTT TTTT 3' "

BASE COUNT 121 a 121 c 179 g 174 t

ORIGIN

alignment_scores: Quality: 199.50 Length: 108
Ratio: 2.934 Gaps: 8
Percent Similarity: 62.963 Percent Identity: 45.370

alignment_block:
US-09-554-547-16 x AA670823 ..

Align seg 1/1 to: AA670823 from: 1 to: 595

```

53 LeuthraLaseRIegLyGlyGluValArgLeuGlyArgAl 69
|||||.....:|||||.....:|||||.....:|||||.....:
9 CTACTCCGAGCGCTGCTACGCTATGCGACAGCTTTGGAGGC...CT 55
69 aGlyValGlyValSerSerTyrr.....GlyTyrrGlyTyrrP 81
111 .....:|||||.....:|||||.....:|||||.....:
56 GGGCTGTGATGTATACAGCATCCGAGCTGGCTGCTGCTGCTATG 105
81 roSerTrpGlyTyrrProTyrrGlyGlyTyrrGlyTyrrGly 97
|||||.....:|||||.....:|||||.....:|||||.....:
106 GAGGCTTCGATATGCG...TCGCTATGAGAGGCTTGGAGGCTTTGGA 152
98 .....GlyTyrrGlyTyrrAspGlnGlyPheGlySerAlaTyrrG 111
|||||.....:|||||.....:|||||.....:|||||.....:
153 TATGCTCTCGCTATGAGAGCTAC.....GGATACGCTCTGCGCTATG 196
111 yGlyTyrr.....ProGlyTyrrGlyTyrrTyrrP 122
|||||.....:|||||.....:|||||.....:|||||.....:
197 AGGCTTGGAGGCTTTGATATGCTCTGCTATGAGAGGCTTCGATATG 246
122 roSerGlyTyrrGlyGly.....GlyTyrrGlyGlySerTyrrGly 134
|||||.....:|||||.....:|||||.....:|||||.....:
247 GCTCTGCGCTATGAGGCTTTGGAGGCTTTGGATATGCTGCTGCTATGGA 296
135 Gly.....SerTyrrGlyGlySerTyrr 141
|||||.....:|||||.....:|||||.....:|||||.....:
297 GGCTATGATATGCTCTGCTTT 320

```

seq_name: gb_est1:AM233247

seq_documentation_block:

LOCUS AM233247 504 bp mRNA EST 30-JUL-2001
DEFINITION f330a09.y1 zebrafish adult olfactory Danio rerio cDNA 5' similar to
contains element M81 repetitive element ;, mRNA sequence.
ACCESSION AM233247
VERSION AM233247.1 GI:6565550
KEYWORDS EST.

SOURCE zebrafish.
ORGANISM Danio rerio
Eukaryota: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi; Cypriniformes; Cyprinidae; Rasbora; Danio.
REFERENCE 1 (bases 1 to 504)
AUTHORS Clark, M., Johnson, S.L., Lehrach, H., Lee, R., Li, F., Marra, M., Eddy, S., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.
TITLE Washu Zebrafish EST Project 1998
JOURNAL Unpublished (1998)
COMMENT Contact: Stephen L. Johnson
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: zbrafish@watson.wustl.edu
CNA Library Preparation: John Ngai CNA Library Arrayed by: Matthew Clark. DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: Genome Systems, St. Louis, Missouri (web address: www.genomesystems.com) (email contact: info@genomesystems.com) and Research Genetics, Huntsville, Alabama (web address: www.resgen.com) (email contact: info@resgen.com) and Ressourcenzentrum Primardatenbank, Berlin, Germany (web address: www.rzpd.de)
Seq primer: T3 ET from Amersham
High quality sequence stop: 440.
FEATURES
source Location/Qualifiers
1..504
/organism="Danio rerio"
/db_xref="taxon:7955"
/clone_lib="zebrafish adult olfactory"
/sex="mixed"
/tissue_type="Olfactory rosettes"
/dev_stage="adult"
/lab_host="D10H (Gibco BRL)"
/note="Vector: pSPORT1; Site 1: NotI; Site 2: SalI; This is a directionally cloned cDNA library from adult zebrafish olfactory epithelium."
BASE COUNT 104 a 120 c 200 g 80 t
ORIGIN
alignment_scores:
Quality: 198.00 Length: 100
Ratio: 3.094 Gaps: 6
Percent Similarity: 64.000 Percent Identity: 48.000
alignment_block:
US-09-554-547-16 x AM233247 ..
Align seg 1/1 to: AM233247 from: 1 to: 504
58 GYGLYGLVAlaIAlaArg...LeuGlyGlyArgAlaGlyValGlyVa 73
||||| |||::||| |||::|||
37 GGTGGAGCCCGCGCGCGAGAGGAGGAGGAGGAGGAGCATGGCGCG 86
73 lserseTyrglyTyrglyTyrglyTyrglyTyrglyTyrglyTy 90
::: |||::||| |||::|||
87 CAATCAAAATGGCTTCGCG...GGCGGTAGAGGAGGAGGAGGCT 124
90 YRGLGLTYRGLTYRGLTYRGLTYRGLTYRGLTYRGLTYRGLTY 104
||||| |||::||| |||::||| |||
125 ACGGGGCGTACGGGGCGGCTATGGCGGAGATACGGCGGAAACGTGGC 174
105 GYpGhlyserAlaTyrglyTyrglyTyrglyTyrglyTyrglyTy 121
||||| |||::||| |||::||| |||
175 GGCTACGGCGGTACGTATGAGAGT...GGAGGCTACGCT... 210
121 rProserGlyTyrglyGlyTyrglyGlyTyrglyGlyTyrgly...G 135

211GGCGGTACGAGAGTACGAGAGGAGGAGGAGGAGCAATGG 256
135 lserTyrglyGlyserTyrglyTyrglyTyrglyTyrglyTyrglyTy 151
||||| |||::||| |||::|||
257 GGGGTATGGCGGTGGAACGTTACAGTACTTTGGCAGCGATACGCG 306
seq_name: gb_est1:A1878240
seq_documentation_block:
LOCUS A1878240 531 bp mRNA EST 07-JUN-2001
DEFINITION IMAGE:3725012 5' similar to SW:R031.XENIA P51968 HETEROGENEOUS NUCLEAR RIBONUCLEOPROTEIN A3 HOMOLOG 1; contains element MER22 repetitive element ;, mRNA sequence.
ACCESSION A1878240
VERSION A1878240.1 GI:5552370
KEYWORDS EST.
SOURCE zebrafish.
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi; Cypriniformes; Cyprinidae; Rasbora; Danio.
REFERENCE 1 (bases 1 to 531)
AUTHORS Clark, M., Johnson, S.L., Lehrach, H., Lee, R., Li, F., Marra, M., Eddy, S., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.
TITLE Washu Zebrafish EST Project 1998
JOURNAL Unpublished (1998)
COMMENT Other ESTs: fc52c11.x1
Contact: Stephen L. Johnson
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: zbrafish@watson.wustl.edu
CNA Library Preparation: Matthew Clark. CNA Library Arrayed by: Matthew Clark. DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: Genome Systems, St. Louis, Missouri (web address: www.genomesystems.com) (email contact: info@genomesystems.com) and Research Genetics, Huntsville, Alabama (web address: www.resgen.com) (email contact: info@resgen.com) and Ressourcenzentrum Primardatenbank, Berlin, Germany (web address: www.rzpd.de)
Seq primer: T3 ET from Amersham
High quality sequence stop: 477.
FEATURES
source Location/Qualifiers
1..531
/organism="Danio rerio"
/db_xref="taxon:7955"
/clone_image="3725012"
/clone_lib="Zebrafish Washu MPMG EST"
/sex="mixed"
/tissue_type="26 somite embryos, adult livers, shield stage embryos"
/lab_host="XLI-blue MRF"
/note="Vector: pSPORT1; Site 1: NotI; Site 2: SalI; 1st strand cDNA was primed with a Not I - oligo(dT)15 primer 15'pGACATGTTCTAGATCGGAGCGCGCGGCTTTTCTTTTCTTTT3'; double-stranded cDNA was ligated to Sal I adaptors (BRL), digested with Not I and cloned into the Not I and Sal I sites of the pSPORT1 vector (BRL). Library was constructed by Matthew Clark (Lehrach Lab; ICRP, London and Max Planck Institut fuer Molekulare Genetik, Berlin). cDNAs for EST analysis were selected following oligonucleotide hybridization fingerprinting of arrayed clones from zebrafish late somitogenesis (26 ss), adult liver or embryonic shield stage (5.6 h) libraries. Fingerprint data were used to computationally cluster cDNAs, and a single cDNA from each cluster was chosen for sequencing.

In some cases multiple members of the same cluster were sequenced to assess clustering parameters or single clones were sequenced additional times to assess quality control."

BASE COUNT 132 a 122 c 200 g 75 t 2 others

ALIGNMENT_SCORES:

Quality: 191.50 Length: 157
Ratio: 2.152 Gaps: 10
Percent Similarity: 56.688 Percent Identity: 40.127

ALIGNMENT_BLOCK:

US-09-554-547-16 x AI878240

Align seg 1/1 to: AI878240 from: 1 to: 531

```

2 LysAlaPhePheValLeuSerLeuSerThrAlaAlaLeuThrAsnAl 18
  ::::::::::: ||| ||| |||
56 GAACGCGTGTTCGCTTGTCCACTTCGAAAGACACAGCCTCTGCAGACA 105
  ::::: ||| ||| |||
18 AAlaArgAlaAlaArgLeuGlySerAspLeuAspThrPheGlyArgValH 35
  ::::: ||| ||| |||
106 AAGCGCTGCTCAAGTTCACATGATCAACGACACAGAGCT..... 148
35 IsGlyAsnLeuTyrAlaGlyIleGluArgAlaGlyProArgGlyTyrPro 51
  ::::: ||| ||| |||
149 .....GGAGCTGACAGAAACCGCTCACGAAACACAGACATCCA 184
52 GlyLeuThrAlaSerIleGlyGlyGluValGlyAlaArgLeuGlyGly 68
  ::::: ||| ||| |||
185 GCGTCGCCGTGGAGCCGCGCGCGGAGCA.....GGAGAGCGGAGGAA 228
68 rGAlaGlyValGlyValSerSer.TyrGlyTyrGlyTyrProSerTrpG 84
  ::::: ||| ||| |||
229 GA...GGCATGGGCGCCCAATCAAAATGCTTCGCGC.....GG 263
84 YTyrProTyrGlyTyrGlyGlyTyrGlyGly.....TyrGlyGlyT 99
  ::::: ||| ||| |||
264 CGGTAAGACGAGAGCGTACGGCGGCTACGGGCGCGCTATGCGGAGGAT 313
99 YrGlyGlyTyrAspGlnGlyPheGlySerAlaTyrGlyGlyTyrProGly 115
  ::::: ||| ||| |||
314 ACGCGCGAAGACGATGGCGGCTACGGCGTATGGAAGT.....GGA 357
116 TTrTyrGlyTyrTyrTyrProSerGlyTyrGlyGlyTyrGlyGlySe 132
  ::::: ||| ||| |||
358 GCGTACGCT.....GCGGTTACGACGCTGATACGGAGAGG 395
132 rTyrGly.....GlySerTyrGlyGlySerTyrThrTyrProAsnV 146
  ::::: ||| ||| |||
396 CTACGCTGACGACATGGCGGCTATGGCGTGAACCGTTACGCTGACT 445
146 AAlaArgAlaSerAlaGly 151
  ::::: ||| |||
446 NTGCGACGCGATACGCGC 462
seq_name: gb_est2:BI306336
seq_documentation_block:
LOCUS BI306336 606 bp mRNA EST 20-JUL-2001
DEFINITION NL_4.D22 Drought stress (leaf) Oryza sativa cDNA clone NL_4.D22 3',
mRNA sequence.
ACCESSION BI306336
VERSION BI306336.1 GI:14981658
KEYWORDS EST.
SOURCE Oryza sativa.
ORGANISM Oryza sativa.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Erihraloidae; Oryzaceae; Oryza.
REFERENCE 1 (bases 1 to 606)

```

AUTHORS

Reddy, A.R., Ramakrishna, W., Chandrasekhar, A., Nagabhushan, I., Ravindrababu, P., and Bennetzen, J.L.

TITLE

Novel EST enrichment with normalized cDNA libraries from drought stressed rice (Oryza sativa L. cv Nagina 22)

JOURNAL

unpublished (2001)

COMMENT

Contact: Reddy AR
Department of Plant Sciences, School of Life Sciences
University of Hyderabad
P.O. Central University, Hyderabad-500 046, A.P., India
Tel: 0091-40-3010265
Fax: 0091-40-3010145
Email: arjuls@uohyd.ernet.in
Insert Length: 606 Std Error: 0.00
Plate: 4 row: D column: 22
Seq primer: GTAAACGACGCGCAGTGT.
Location/Qualifiers

FEATURES

SOURCE

1. 606
/organism="Oryza sativa"
/cultivar="Nagina 22 (Indica sub sp)"
/db_xref="taxon:4530"
/clone="NL_4.D22"
/clone_lib="Drought stress (leaf)"
/tissue_type="Entire leaf tissue"
/dev_stage="35 day-old seedlings"
/note="Organ: Leaf; Vector: T73Pac; ESTs from normalized leaf cDNA library from drought stressed seedlings"

BASE COUNT

129 a 218 c 146 g 113 t

ALIGNMENT_SCORES:

Quality: 191.00 Length: 134
Ratio: 2.418 Gaps: 6
Percent Similarity: 58.955 Percent Identity: 41.791

ALIGNMENT_BLOCK:

US-09-554-547-16 x BI306336/rev

Align seg 1/1 to reverse of: BI306336 from: 1 to: 606

```

11 SerThrAlaAlaLeuThrAsnAlaAlaArgAlaGlyArgLeuGlySerS 27
  ::::: ||| ||| |||
585 AGTCCAGATCATCAACGATCGGGAACGCGGAGGTCTCGCGCTTCGCG 536
27 pIeu AspThrPheGlyArgVal..HisGlyAsnLeuTyrAlaGlyLe 42
  ::::: ||| ||| |||
535 TTGTGACGTTTTCGAGCGAGCGCATGCGCGGACCATGCAATCGAGGAT 486
43 GluArgAlaGlyProArgGly.TyrProGlyLeuThrAlaSerIleGly 59
  ::::: ||| ||| |||
485 GAACGG.....CAAGAGCTTGACGGCGCGCCACATCACCCTTAACG 445
59 IyGluValAlaAlaArgLeuGlyGlyArgAlaGlyValGlyValSerSer 75
  ::::: ||| ||| |||
444 AGGCCGAGTCCGCCCGCTCCGAGCGGAGCGCGC..... 410
76 TTrGlyTyrGlyTyrProSerTrpGlyTyrProTyrGlyGlyTyrGly 92
  ::::: ||| ||| |||
409 GGTGCTACGGCCAGCCCGCGGAGCGGAGCGGAGCTTACGGGTGCG...GGCG 363
92 YTrGlyGlyTyrGlyGlyTyrGlyGlyTyrAspGlnGlyPheGlySerA 109
  ::::: ||| ||| |||
362 CTACGGTGTGTGGGCGCGCGGTGCTACGGCCACAGCCCGCGAAGCGG 313
109 lATrGlyGlyTyrProGlyTyrTyrGlyTyrTyrTyrProSerGlyTyr 125
  ::::: ||| ||| |||
312 GCTACGCGGTGGCGGCTGCTACGGTGGCGCGCGTGGTGGCGGCTAC 263
126 GylGlyGlyTyrGlySerTyrGlyGlySerTyrGlyGlySerTyrGly 141
  ::::: ||| ||| |||
262 GTGTGGCTACGGCAACCGCGCGCGGCGGCACTCCGACGGAGACTGG 215
seq_name: gb_est1:AI795569

```

```

11 ||||| ||||| ::|||
276 GGCCCCAGGTGTTATGGC.....ACGGCGATTACGCAGTAT 316
127 GYGTYrGYrGLySeTyrgLySerTyrgLySeTyrgLy.....GlySeTy 141
11111 |:::~| ||||| ~||
317 GCGCGAGCAGTAGTACTGTGTGGGACACAGTGCTCGGGCTTAG 366
141 rThrTy 143
111
367 CAGGTAT 373

seq_name: gb_escI:AW000266

seq_documentation_block:
LOCUS AW000266 649 bp mRNA EST 08-SEP-1999
DEFINITION G14009H04.y1 614 - root cDNA library from Walbot Lab Zea mays CDNA
mRNA sequence.
ACCESSION AW000266
VERSION AM000266.1 GI:5847187
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays.
Eukaryote: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PAC
clade: Panicoidae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 649)
AUTHORS Walbot,V.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford
UNIVERSITY
JOURNAL Unpublished (1999)
COMMENT Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
FEATURES
source
Location/Qualifiers
1..649
/oranism="Zea mays"
/cultivar="W23"
/db_xref="taxon:4577"
/clone_lib="614 - root cDNA library from Walbot Lab"
/lssue_type="root"
/dev_stage="3-4 days old"
/lab_host="XLR"
/note="Organ: root; Vector: pBluescriptII SK+; Site:1: EcoRI; Site_2: XhoI; 3-4 days old root tissue from Walbot Lab (Lw)"
BASE COUNT 131 a 115 c 220 g 182 t 1 others
ORIGIN
Alignment_scores:
Quality: 191.00 Length: 119
Ratio: 2.768 Gaps: 8
Percent Similarity: 57.983 Percent Identity: 47.059
alignment_block:
US-09-554-547-16 x AW000266 ..
Align seg 1/1 to: AW000266 from: 1 to: 649
39 TTAAGlAGlYleGlUArgAlAGlProAAGlYTr.....ProgI 52
||||:|||| |:::~| ||||| ~||
40 TATGGGGGCTTATCGATCTCGCAGCAGGTTATGATGATCATGCTG 89
52 yLeuThralaserilegLYglYlwaLYglYlaatrgLaegLYglYARG 69
1 |:::~| ||||| ~|| |:::~| ||||| ~||
90 CGCAGAGCAGCATATGCTATGCTAGAAGGATGCTATGAGGCCAAAT 139
69 laglYalValcIvalserSerTyrgLyTrYrProSer..... 82

```

```
140 CTGGCTTTGGG...TCTAGTTTGGTGGCTTATGCGATCCATCAT 186
83 TrpGlyTyrProTyrGlyGlyTyrGlyTyrGly.....TyrGly 97
187 CGAGCTGCTGCCCTATGTCATATGTCCTATGTCGTGGCGCCTATGG 236
97 yGly.....TyrGlyGlyTyrAspGlnGlyPheGlySerAlaTyrGlyG 112
237 AGGAGGTGCCCTACGAGCGCGCTGCTATGAGGTGGTGGCTGCTATGGTG 286
112 LyTyrProGlyTyrTyrGlyTyrTyrTyrProSerGlyTyrGly..... 126
287 CCCCCCAGCGCTTATGCC.....ACTGGCGGATACGGCAGTTAT 327
127 GlyGlyTyrGlyGlySerTyrGlyGlySerTyrGly.....GlySerTyr 141
328 GCGCGACGAGTATGTCGTGGGAGCACAGGTGCTCGGGGTTCTAG 377
141 rThrTyr 143
378 CAGGTAT 384
```

THIS PAGE BLANK (USPTO)

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☒ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☒ **FADED TEXT OR DRAWING**
- ☒ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☐ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.

THIS PAGE BLANK (USPTO)